Official Title: A Phase III, Multicentre, Randomized, Parallel-Group, Double Blinded, Placebo Controlled Study to Evaluate the Efficacy and Safety of Ocrelizumab in Adults With Primary Progressive Multiple Sclerosis

NCT Number: NCT01194570

Document Dates: SAP Version 5: 20-March-2023

STATISTICAL ANALYSIS PLAN

TITLE: A PHASE III, MULTICENTRE, RANDOMIZED, PARALLEL-GROUP, DOUBLE BLINDED, PLACEBO CONTROLLED STUDY TO EVALUATE THE EFFICACY AND SAFETY OF OCRELIZUMAB IN ADULTS WITH PRIMARY PROGRESSIVE MULTIPLE SCLEROSIS

PROTOCOL NUMBER: WA25046

STUDY DRUG: Ocrelizumab

VERSION NUMBER: 5

IND NUMBER: 100,593

EUDRACT NUMBER: 2010-020338-25

SPONSOR: F. Hoffmann-La Roche Ltd

PLAN PREPARED BY: [redacted] Ph.D.
[redacted] M.Sc.

DATE FINAL: 20 December 2013

DATES AMENDED: Version 2: 30 October 2014
Version 3: 28 January 2015
Version 4: 03 September 2015
Version 5: See electronic date stamp on the last page of this document

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Version 1 (20 December 2013)

Based on Version D of the Protocol WA25046 (15 June 2012).

Version 2 (30 October 2014, Protocol Version D)

Updates to Version 1 made at this time include the following substantive changes:

- The schedule of assessment was removed to reduce redundancy with the protocol.
- More details on the analyses timing were added, and a maximum duration of the blinded treatment period was introduced following re-estimation of the trial duration based on actual event rates. These changes were consistent with the protocol, and intended to better define the likely end of trial.
  - If the prediction does not ensure at least 253 confirmed disability progression (CDP) events at Week 120, this procedure will be repeated 6 weeks later.
  - The primary database lock will occur no later than three years after the last patient was randomized, even if the 253 events have not been reached at this timepoint.
- Clarifying that the safety population includes all patients dosed to adhere to the standard definition.
- The intent-to-treat (ITT) population includes all patients randomized to adhere to the standard definition.
- The secondary efficacy outcome “Change in Total Volume of T2 Lesions on MRI Scans of the Brain from Baseline to Week 120” was moved down from the third to the fourth rank in the testing hierarchy. Consequently, the “Change in 25-foot timed walk” moved from fourth to the third rank in the testing hierarchy.
- “Percent Change on MRI from Week 24 to Week 120 in Total Brain Volume” and “Change from Baseline in Quality of Life as Measured by the SF-36 Physical Component Score (PCS)” were added at the end of the list of secondary endpoints.
- EMA Scientific Advice was implemented to increase statistical rigor (e.g., replacing Van Elteren Tests and last-observation carried forward [LOCF] method) with Mixed-Effect Model Repeated Measures [MMRM] analysis).
- The “Change in Total Volume of T2 Lesions on MRI Scans of the Brain from Baseline to Week 120” will be analyzed as “Percent Change in Total Volume of T2 Lesions on MRI Scans of the Brain from Baseline to Week 120” to align with the other MRI measures (brain volume, white and grey matter).
- The following exploratory endpoints were added:
  - The change in fatigue, as measured by the Modified Fatigue Impact Scale (MFIS) total score and subscale scores (Physical impact, Cognitive impact, and Psychological impact) from baseline to Week 120.

– The change in quality of life, as measured by the SF-36v2 Mental Component Summary (MCS) Score from baseline to Week 120.

– Time to CDP over the treatment period, defined as an increase in Expanded Disability Status Scale (EDSS) that is sustained for at least 12 weeks (0.5 or 1, same criteria as for the primary endpoint time to 12-week CDP) or a 20% increase in 25-foot timed walk that is sustained for at least 12 weeks or a 20% increase in 9-hole peg test that is sustained for at least 12 weeks.

- Additional sensitivity analysis:
  - For the time to onset of CDP for $\geq$24 weeks, the influence of early progression events on treatment effect will also be evaluated by omitting the EDSS assessments performed between randomization and the Week 12 visit ($\leq$83 days after randomization) with use of the ITT population.
- Adverse events:

– Detailed definition of serious adverse events (SAEs) was added; that is, “SAEs will be defined as all SAEs including serious MS relapses and serious IRRs.”

– The derivation of the 95% CI for the number of adverse events (AEs) per 100 patient-years was aligned with the Phase III relapsing multiple sclerosis (RMS) studies.

Version 3 (28 January 2015, Protocol Version D)

Updates made at this time include the following substantive changes:

- Change of the baseline definition for assessments other than EDSS to include the baseline/Day 1 visit into the baseline. Blinded review of the data indicated that baseline assessment on Day 1 occurred after the day of randomization in many cases. Thus, restricting the baseline to up to and including the day of randomization would have led to a considerable number of patients without baseline assessments.
- The change in 25-foot timed walk from baseline to Week 120 was added to the subgroup analyses.

Version 4 (03 September 2015, Protocol Version E)

This Statistical Analysis Plan has been updated to align with the amended study protocol (WA25046, Version E, 25 February 2015). Updates made at this time include the following substantive changes. The rationale for the update is to provide full and complete information on these important aspects prior to database lock.

Changes to the statistical analyses methods:

- Experience from the RMS Phase III program and blinded review of WA25046 data indicated that the use of MMRM on the absolute change in 25-foot timed walk from baseline to Week 120 would violate the assumptions of normal distribution of the residuals. Therefore, we will analyze relative change from baseline. The p-value of the non-parametric ranked analysis of covariance (ranked ANCOVA) on the percent

change from baseline will be reported. The LOCF method will be used to impute missing values. Assuming a higher withdrawal rate in the control than in the treatment arm and a monotonous increase in the 25-foot timed walk, this method is expected to be conservative. Since the ranked ANCOVA does not provide estimates of the treatment effects, these will be derived using MMRM analyses on the log-transformed post-baseline/baseline values.

In addition, rules for handling outliers in 25-foot timed walk were added based on the Multiple Sclerosis Functional Composite (MSFC) Manual and published data on reference ranges from healthy subject cohorts.

- Experience from the RMS Phase III program (WA21092 and WA21093) and blinded review of WA25046 data indicated that for the percentage change in T2 lesion volume, the MMRM analyses would violate the assumptions of normal distribution of the residuals. Therefore, the same analysis approach (ranked ANCOVA, MMRM) as for the change in 25-Foot Timed Walk will be used, and the title was changed from percentage change to change in T2 lesion volume.
- Additional sensitivity analyses: Based on the results of the Phase III RMS studies (WA21092, WA21093) and the blinded review of WA25046 data, a sensitivity analysis adjusting for presence of Gadolinium-enhancing lesions and baseline EDSS (≤ 5.5 vs > 5.5) at baseline was added. Blinded review of WA25046 data indicated that there are patients with relapses in the study. Since primary progressive multiple sclerosis (PPMS) is characterized by a progressive course from disease onset typically without superimposed discrete clinical attacks or relapses (Ebers 2004), the analyses of time to CDP for at least 12 weeks and the time to CDP for at least 24 weeks will be performed excluding patients with clinical relapses.
- Body weight and body mass index have been added to the subgroup analyses.

Other changes:

- The standard operating procedure describing the checks done on EDSS values has been added, which is effective since January 2013, to provide further detail on the EDSS cleaning algorithm.
- Consistent with the protocol, detailed reasons for patient exclusion from per-protocol population have been added.
- The definition of disease-modifying treatments (DMTs) was updated to account for most recent available therapies.
- The definition of protocol-defined relapses has been updated to align with the current protocol.
- SAS codes have been included for statistical models for primary and secondary analyses.
- Detailed updates have been made to the safety analysis section, including the definition of the duration of observation, details on the derivation of AEs and SAEs and the assignment of AEs to a specific dose, the detailed description of the

analyses done for selected AEs (i.e., infusion-related reaction [IRRs], infections, opportunistic infections [OIs], malignancies).

- To have consistent description of safety analysis across the RMS and progressive multiple sclerosis trials and integrated safety analyses, the details of the safety analyses were added in this document.

Additional minor changes have been made to each version to improve clarity and consistency.

Version 5 (March 2023, Protocol Version F, G, H, I, J and K)

This SAP is updated to describe the analyses to be performed on data from the extended controlled period and open-label extension (OLE) period. As additional data collection and objectives may differ from double-blind phase, SAP requires modification. Moreover, several modifications were made based on the availability of the data. A list of these modifications can be found below. A table is also added to display the status of each reporting event.

The following study periods will be considered for the final analyses:

- Efficacy analyses will be performed for the extended controlled treatment period (ECP), i.e., from randomization up to the first dose of Ocrelizumab in OLE.
- Efficacy analyses will be performed for the OLE period, i.e., from the first dose of Ocrelizumab in OLE.
- Efficacy analyses will be performed for the combined periods of ECP and OLE.
- Safety analysis will be performed separately for the OLE and OLE safety follow-up (SFU) periods as well as for all study periods combined.

Additional endpoints for the extension periods:

- Time to confirmed EDSS score $\geq$7.0 (time to requiring a wheelchair) for at least 24 weeks during the combined ECP and OLE periods.
- Time to confirmed EDSS score $\geq$7.0 (time to requiring a wheelchair) for at least 48 weeks during the combined ECP and OLE periods.
- Change from baseline in total non-enhancing T1 lesion volume.
- Annualized change in MRI parameters including brain volumes, T1 lesion volume and T2 lesions volume.

Protocol Updates:

Due to various protocol amendments, SAP is updated to reflect changes or updates made in the protocol amendment. Below are some of such changes;

- Pharmacokinetic (PK)/human anti-human antibody (HAHA) collection has been stopped since Dec 2019.

- Ocrelizumab infusion must be suspended in the event of an active TB infection or if a female patient is pregnant or breastfeeding in OLE period. Ocrelizumab infusions may be restarted at the discretion of the Investigator and based on individual benefit-risk assessments, but only upon resolution of the active TB infection or after completion of pregnancy and breastfeeding.
- The requirement for continued B cell monitoring for participants whose B cells are not repleted (i.e., returned to baseline levels or the lower limit of normal, whichever is lower) at the end of the SFU period has been removed.
- Patients starting commercial ocrelizumab or other DMTs will discontinue from the study completely and will not enter SFU.
- The OLE treatment phase is extended till 31 Dec 2022 to allow additional long-term efficacy and safety data. An option is introduced for all ongoing participants of Study WA25046 to enroll into a new OLE study (MN43964) prior to or following the closure of Study WA25046, latest by End of 2022.

Additional minor changes have been made to improve clarity and consistency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

1. BACKGROUND

Study WA25046 is a Phase III, multicenter, randomized, parallel group, double blind, placebo controlled study to evaluate the efficacy and safety of ocrelizumab in adults with primary progressive multiple sclerosis (PPMS).

To date, no treatment has been demonstrated to significantly slow the progression of disability in patients with PPMS. Currently, PPMS remains a severely disabling condition with very high unmet medical need.

Ocrelizumab (RO4964913) is a humanized, glycosylated monoclonal antibody directed against the CD20 antigen present on B cells. In addition to this Phase III clinical study of ocrelizumab in PPMS (Study WA25046), the safety and efficacy of ocrelizumab in multiple sclerosis (MS) is also being evaluated in two ongoing randomized, controlled Phase III clinical trials (WA21092 Opera I and WA21093 Opera II) and one Phase II study in relapsing-remitting MS (RRMS) (WA21493).

The Phase II study (Study WA21493) in RRMS, currently in the open-label extension (OLE) stage and unblinded for efficacy analysis in October 2009, showed a rapid effect on B-cell depletion, magnetic resonance imaging (MRI) findings, and clinical activity. After 24 weeks of treatment, both doses of ocrelizumab (2×300 mg and 2×1000 mg) demonstrated a strong effect with a statistically significant reduction in signs of disease activity as measured by MRI lesions (the primary endpoint) relative to placebo. Both ocrelizumab dose regimens showed statistically significant differences in the total number of gadolinium-enhancing lesions at Weeks 12, 16, 20, and 24. The relative reductions in the number of lesions compared with placebo were 89% and 96% (for 2×300mg and 2×1000mg doses of ocrelizumab, respectively). Both ocrelizumab doses showed a statistically significant reduction in the annualized relapse rate (ARR) compared with placebo. Moreover, both ocrelizumab doses were superior to the active comparator, Avonex®, for the MRI primary endpoint and ARR.

The Phase III relapsing multiple sclerosis (RMS) (WA21092, WA21093) studies were locked in May/June 2015, and the clinical study reports (CSRs) are being written.

On 6 March 2013, the United States (U.S.) Food and Drug Administration (FDA) granted Fast Track designation to the investigation of ocrelizumab for PPMS.

The study has been unblinded at the end of the blinded treatment period in September 2015 (clinical cutoff date [CCOD] 24 July 2015) when reaching at least 120 weeks after the last patient randomized provided that the planned number of confirmed disability progression (CDP) has reached 253, (number of events included in the primary analysis: 256). After the positive readout, patients were given access to the OLE starting from 12 October 2015. In the period between the CCOD for the Primary CSR and OLE initiation, the Sponsor, sites, and patients were sequentially unblinded. This period corresponds to the Extended Controlled treatment period (ECP). The ECP

encompasses all efficacy data from the double-blind period (DBP) (previously reported in the Primary CSR) plus any additional efficacy data collected during the controlled treatment period up to the time at which the patient received their first open-label dose of ocrelizumab or until withdrawal.

This document has been updated (Version 5) to describe the further analyses to be performed, on:

1. the ECP, including additional data collected after the primary CCOD up to the entry into OLE.
2. the OLE period, including additional data collected from the first dose of Ocrelizumab in OLE.
3. the combined periods of ECP and OLE.

2. STUDY DESIGN

Study WA25046 will evaluate the safety and efficacy of a minimum of five treatment doses of two 300 mg intravenous (IV) infusions of ocrelizumab separated by 14 days, occurring at a scheduled interval of 24 weeks, compared with placebo in adults with PPMS. Patients will be treated for a minimum of 120 weeks, representing at least five 24 week treatment doses, and either when approximately 253 CDP events have occurred or the last patient randomized has been treated for 3 years at the latest.

This study consists of the following study periods: a screening period, a blinded treatment period, an open label treatment period, and a safety follow-up (SFU) period.

The primary analysis of the study will be performed when the last patient enrolled has been treated for at least 120 weeks. All patients will stay on their randomized treatment until the primary efficacy analyses. In addition, in the event that the number of disability progression events is lower than anticipated at the time the last patient reaches Week 120, the blinded treatment period may be extended by additional treatment doses for all patients, to maintain statistical power to detect a treatment difference, until approximately 253 confirmed CDP events have occurred (see Section 2.3).

The schedule of assessments are included in the protocol.

This document describes the primary analysis from double-blind phase, additional analysis from extended controlled phase and OLE phase. The Sponsor has now made the decision not to extend this study beyond 31 December 2022. Instead, a rollover extension study (MN43964) is set up to ensure that participants of Study WA25046 (together with participants from other Parent studies) can continue their ocrelizumab treatment or SFU as applicable without interruption and allowing for valuable long-term data to continue to be collected.

Screening: 4 weeks of screening period to evaluate patient eligibility.

Blinded phase: Patients going through at least 120 weeks of study treatment representing 5 cycles, each of 24 weeks. Study will be unblinded when the last enrolled patient completes at least 120 weeks, and the total number of CDP is approximately 253, whichever occurs later.

Extended Controlled treatment Period (ECP): The ECP will include the DBP plus additional follow-up prior to OLE enrollment or treatment withdrawal. Patients from the DBP continued in their randomized treatment group for some additional time until sequentially switching into the OLE. This time period between the end of the DBP and the start of the OLE includes: approximately 3 additional months of blinded controlled treatment and approximately 6 additional months of controlled follow-up during which time patients will gradually be unblinded and switched into the OLE.

Open-Label Extension(OLE) phase: Eligible patients will continue to receive additional treatment cycles of open label ocrelizumab following the ECP. The OLE Phase will continue as per local regulation or should the Sponsor decide to terminate the ocrelizumab program for MS. Unless terminated earlier for any of the reasons mentioned above, all patients will continue their treatment with open-label ocrelizumab as per the protocol until an anticipated study completion date 31 December 2022. All patients should complete the trial and migrate to the new roll-over extension study (MN43964) on or before 31 December 2022.

Safety follow-up period: Patients who discontinue treatment for any reason during the following periods will be entered into the SFU Period:

- During or after completion of the blinded treatment period
- During the OLE Phase Screening Period
- During the OLE Phase
- Patients who choose not to enter the OLE Phase or are not eligible for the OLE Phase after blinded treatment period

Patients who start treatment with commercial ocrelizumab or other disease-modifying treatments (DMTs) will discontinue from the study completely and will not enter or continue in the SFU Period. Patients in SFU will be followed for 48 weeks or until the anticipated study completion date of 31 December 2022. All patients should complete the trial and migrate to the new roll-over extension study (MN43964) on or before 31 December 2022.

2.1 OUTCOME MEASURES

2.1.1 Efficacy Outcome Measures

The primary efficacy endpoint is the time to onset of 12-week CDP during the blinded treatment period. The blinded treatment period is not fixed but will depend on the timing of the primary analysis (Section 2.3).

Disability progression is defined as an increase of ≥1.0 point from the baseline Expanded Disability Status Scale (EDSS) score if the baseline EDSS score is ≤5.5 points or an increase of at least 0.5 points if the baseline EDSS score is >5.5 points–for which change is not attributable to another etiology (e.g., fever, concurrent illness, MS relapse or exacerbation, or concomitant medication).

The EDSS will be assessed in all patients by an independent examining investigator at screening, baseline, and every 12 weeks (at a regularly scheduled visit) during the blinded treatment period of the study, during the SFU period, at any unscheduled visits, and at withdrawal from treatment and end of study visits. Additional EDSS assessments for individual patients may be requested between visits (i.e., during an MS relapse). The baseline EDSS score is calculated as the average of the assessments at screening and the Day 1 visit (for further details, see Section 4.5.1).

Confirmation of disability progression must occur at a regularly scheduled visit that is ≥12 weeks (≥84 days) after the initial disability progression. Initial disability progression can occur at a scheduled or unscheduled visit after the date of randomization. The EDSS assessments at unscheduled and non-confirmatory scheduled visits (if any) between the initial disability progression and the confirmation of disability progression should have values equal to or exceeding the minimum change required for progression (e.g., a subject with a baseline EDSS value of 3 must have an EDSS value of ≥4 at all visits between the visit with initial disability progression and the scheduled visits to confirm 12 week or 24 week CDP).

The secondary and exploratory outcome measures to be analyzed are described in Section 4.5.2 and Section 4.5.3, respectively.

2.1.2 Pharmacokinetic Outcome Measures

Ocrelizumab serum concentration time data will be modeled using a population approach. The primary population pharmacokinetic (PK) parameters for ocrelizumab (clearances and volumes) will be estimated by NONMEM analysis of the sparse PK data. Clearances with associated inter-patient variability may be characterized by a saturable and nonsaturable clearance, as well as an inter-compartmental clearance, depending on the final structural model. Volumes with associated inter-patient variability may be characterized by central and peripheral volumes depending on the final structural model. Exposure to ocrelizumab (area under the concentration time curve [AUC]) will be estimated. The selection of other parameters will depend on the final PK model used for this analysis.

2.1.3 Safety Outcome Measures

Safety will be assessed through regular neurologic and physical examinations, vital signs, ECG, and the occurrence of adverse events (AEs). In addition, the following laboratory data will be examined:

- Complete standard hematology, chemistry, and urinalyses assessment
- Circulating B-cell total and subsets, T cells, natural killer cells, and other leukocytes
- Plasma immunoglobulins
- Antibody titers to common antigens
- Anti-drug antibodies (ADAs), previously called human anti-human antibodies (HAHAs)
- HBV DNA in patient core HBV antibody positive at screening

2.2 DETERMINATION OF SAMPLE SIZE

The sample size for the CDP (12 week confirmation) was estimated on the basis of data from a previous rituximab Phase II/III trial in adults with PPMS (Study U2786g, Hawker et al. 2009). For the current study, the 2 year progression rate in patients receiving ocrelizumab is predicted to be 30%, compared with 43% in patients receiving placebo (hazard ratio=0.635). Assumptions of equal exponential survival, exponential dropout, and a log rank test were used to determine the sample size for the time to onset of CDP at 12 weeks. In addition, the following assumptions were made for the study conduct: a 2:1 randomization ratio between the ocrelizumab and placebo arms, a 1 year accrual period with a 3.5 year maximum blinded treatment period, and a dropout rate of 20% over 2 years. On the basis of these assumptions, the predicted power and number of events for the originally planned sample size (N=630) were calculated. These results are presented in Table 1.

Table 1 Predicted Power and Number of Events

Sample Size	Planned
N	630
Type 1 error rate	0.01
Power (%)	80
No. of expected events	253

For the primary efficacy analysis, the null hypothesis will be tested at an alpha level of 0.05 with an associated power of approximately 90%. This will also provide sufficient power (80%) for an alpha=0.01 which may reflect a strong treatment effect.

Due to a late screening boost, the actual number of patients enrolled was 732. If the even rate is as assumed in the protocol, we expect 295 events at study unblinding if

performed 120 weeks after the last patient randomized. This will result in a power of 87% at alpha 0.01, and 96% at alpha 0.05.

Patients may show lower event rates than assumed in the sample size calculations. Therefore, at 120 weeks after the last patient randomized the initially assumed 253 events may not be reached. In this case, the study unblinding will be later than 120 weeks after the last patient randomized. The unblinding will still be triggered by reaching 253 events, and the power remains as originally planned.

If the 253 events are not reached three years after the last patient has been randomized, the study will be unblinded with fewer than 253 events. In this case, the power would be lower than predicted.

2.3 ANALYSIS TIMING

The study will be unblinded when the last enrolled patient completes at least 120 weeks of study treatment, provided that the total number of CDP events for the primary efficacy analysis is approximately 253. This value is based on the Sponsor's best estimation after the last patient finishes the Week 108 visit; that is, when the last patient has reached the Week 108 visit, the number of events at Week 120 (12 weeks later) will be predicted on the basis of ongoing event tracking. At this time, the date of database lock will be fixed if the projected number of events exceeds 253 when the last patient reaches the Week 120 visit, irrespective of whether the 253 events would have been reached 12 weeks later. This procedure should ensure objectivity in fixing the database lock time and adequate precision in reaching the approximate required number of events. If the prediction does not ensure at least 253 events at Week 120, this procedure will be repeated 6 weeks later.

All patients will be unblinded at the time of the primary efficacy and safety analysis. Once approximately 253 events have been reached, the treatment assignments will be unblinded to the Sponsor for the primary analysis and the blinded treatment period will end. The primary analysis will include all data available at the time of the database lock (called primary database lock in the remainder of the document) for both safety and efficacy endpoints, unless stated otherwise.

The primary and secondary efficacy endpoints related to CDP will use all available data in the database at the time of primary database lock. All other endpoints with respect to a specific timepoint (e.g., change from baseline to Week 120) will use data collected for each patient only up to and including the pre-specified timepoint. Any data collected for patients beyond that timepoint will not be included in the endpoint derivation or analysis.

The primary database lock will occur no later than three years after the last patient was randomized, even if the 253 events have not been reached at this timepoint. At the end of the study an additional analysis comprising all data collected including the OLE will be performed.

After the primary database lock, additional data is collected from the ECP, the OLE period and SFU after withdrawing from treatment during any study phase i.e., double-blind or extended treatment period or OLE treatment.

- The extended controlled treatment includes all data from randomization until the day of withdrawal from study treatment or the day before the first dose of ocrelizumab in the OLE period of the study, whichever is earlier.
- The OLE period includes all data from the day of the first dose of ocrelizumab treatment until the day of study completion or withdrawal from study treatment during OLE.
- The combined ECP and OLE period includes all data from randomization until the day of study completion or withdrawal from the study treatment.

The following reporting events are done / planned for this study:

Report	CCOD	Status
Primary analysis CSR	24 July 2015	Completed
3 Months safety update/ Efficacy Data memo	20 January 2016	Completed
Interim CSR (Time to milestone analysis)	03 January 2020	Completed
Pooled safety analyses for safety monitoring and publication of updated interim results	Yearly datacut has been performed since Feb 2017 to conduct interim analyses used for publication purposes	Completed until 2021 and planned for subsequent years until the end of the study
Final CSR	After completion of the ECP, OLE phase or SFU phase of the study (31 Dec 2022)	Planned

CCOD = clinical cutoff date ;CSR = clinical study report, ECP = extended controlled treatment period; OLE = open-label extension; SFU = safety follow-up.

3. STUDY CONDUCT

3.1 RANDOMIZATION

Approximately 630 patients aged 18−55 years inclusive were planned to be randomized in a 2:1 ratio to two groups. An independent Interactive Voice and Web Response System (IxRS) provider conducted the randomization and holds the treatment assignment code. Patients were stratified by region (U.S. vs. rest of world [ROW]) and age (≤45 vs. >45 years).

The patient randomization list was generated by the IxRS with use of a predefined randomization specification. The randomization list is not available at the study center, to the Roche monitors, or to anyone on the Sponsor's side involved in the project. Unblinding of the ocrelizumab dose for individual patients should not occur except in the case of emergency situations. Any request from the investigator for information about

the treatment administered to study patients for another purpose must be discussed with Roche. Unblinding will be performed by means of an IxRS. As per regulatory reporting requirements, Roche will unblind the identity of the study drug for all unexpected serious adverse events (SAEs) that are considered by the investigator to be related to study drug according to safety reference documents (e.g., Investigator's Brochure, Core Data Sheet, and Summary of Product Characteristics). Details of the patients whose data are unblinded during the study will be included in the CSR.

After the end of randomization, the data entered into the IxRS system will be reconciled with the data entered into the Case Report Form (CRF); in particular, the randomization dates and stratification factors (age and region) will be checked. A summary of the discrepancies between the IxRS and CRF (age group and region) will be provided.

For Lithuania and Germany, the year of birth (but not the day and month) is being collected to comply with local regulations. The date of birth will be used as defined in the IxRS specifications.

Because of a late screening boost, 732 patients (instead of the planned 630 patients) were enrolled in the trial. To assess the effect of the increased enrollment numbers on the primary endpoint, a sensitivity analysis will be performed (see Section 4.5.4).

3.2 INDEPENDENT REVIEW FACILITY

The two secondary efficacy endpoints are the change in total volume of T2 lesions on MRI scans from baseline to 120 weeks and the percent change in total brain volume from Week 24 to Week 120. MRI scans will be read by a centralized reading center. The centralized reading center is blinded to the treatment assignment, and the reading is performed in the absence of clinical information. Further details of scanning acquisition sequences, methods, the handling and transmission of the scans, the certification of the site's MRI radiologist and technicians, and the procedures for blinded analysis of the scans at the central reading center are described in a separate MRI technical manual.

3.3 EXPANDED DISABILITY STATUS SCALE CLEANING PROCESS

The primary efficacy endpoint will be derived from the EDSS values recorded at any visits. EDSS assessments are performed by an examining investigator (not the treating investigator), entered into an electronic device, and transferred to a central database for data-cleaning activities. All EDSS results will then be checked in accordance with the standard operating procedure entitled "EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093" (see Appendix 1).

3.4 DATA MONITORING

An independent Data Monitoring Committee (iDMC) will review cumulative data from the trial at approximately 4 month intervals until the primary analysis is complete. After the

primary unblinding, the iDMC will stop reviewing the data. The iDMC will primarily review safety data. No iDMC review is conducted post unblinding of the study.

4. STATISTICAL METHODS

4.1 ANALYSIS POPULATIONS

One patient population will be defined for the purpose of the safety analysis, and two populations will be defined for the efficacy analysis. All efficacy endpoints will be analyzed using the intent-to-treat (ITT) population. The per-protocol population will be used for the primary efficacy endpoint only, to evaluate the influence of major protocol deviations and as a sensitivity check for the ITT analysis.

4.1.1 Intent-to-Treat Population

All randomized patients will be included in the ITT population. Patients who prematurely withdraw from the study for any reason or for whom an assessment is not performed for any reason will still be included in the ITT analysis. Patients who receive an incorrect therapy (different from that which is intended) will be summarized according to their randomized treatment.

4.1.2 OLE ITT Population

All patients receiving at least one dose of OLE as part of the OLE will be included in the OLE ITT population. Patients will be summarized according to their randomized treatment.

4.1.3 Per-Protocol Population

The per protocol population will include all patients in the ITT population who adhere to the protocol and will be summarized according to the randomization arm. Patients may be excluded if they violate the inclusion or exclusion criteria or deviate from the study plan. Specific reasons for exclusion were agreed on based on the final version of the protocol prior to the unblinding of the treatment groups and are documented below. Only those patients with deviations that are deemed to potentially affect the efficacy of study treatment will be excluded from the per protocol population.

The following patients will also be excluded from the per-protocol population:

- Patients with a diagnosis of relapsing-remitting, secondary progressive, or progressive relapsing MS
- Patients with a significant, uncontrolled disease, such as cardiovascular (including cardiac arrhythmia), pulmonary (including obstructive pulmonary disease), renal, hepatic, endocrine, or gastrointestinal or any other significant disease that may preclude patients from participating in the study
- Patients with any concomitant disease that may require chronic treatment with systemic corticosteroids or immunosuppressants during the course of the study

- Patients with a history or with currently active primary or secondary immunodeficiency
- Patients who received previous treatment with B-cell targeted therapies
- Patients who received any previous treatment with alemtuzumab, antiCD4, cladribine, cyclophosphamide, mitoxantrone, azathioprine, mycophenolate mofetil, cyclosporine, or methotrexate or those who had total body irradiation or bone marrow transplantation
- Patients who received any previous treatment with lymphocyte trafficking blockers (e.g., natalizumab, FTY720)
- Patients who received treatment with β-interferons, glatiramer acetate, IV immunoglobulin, plasmapheresis, or other immunomodulatory therapies within 12 weeks prior to randomization
- Patients with known presence of neurological disorders
- Patients who are pregnant or lactating at entry of the study
- Patients who received treatment with any investigational agent within 24 weeks of screening or five half-lives of the investigational drug (whichever is longer) or those who received treatment with experimental procedure for MS (e.g., treatment for chronic cerebrospinal venous insufficiency)
- Patients with no diagnosis of PPMS, in accordance with the revised McDonald criteria (2005; see Protocol, Appendix 1)
- Patients with an EDSS score outside 3−6.5 points at screening
- Patients with a score of <2.0 on the Functional Systems (FS) for the pyramidal system that is due to lower extremity findings at screening
- Patients with a disease duration of >15 years from the onset of MS symptoms in those with an EDSS >5.0 at screening or those with a disease duration of >10 years from the onset of MS symptoms in those with an EDSS at screening ≤ 5.0
- Patients who received no dose of ocrelizumab/ocrelizumab placebo
- Patients who received ocrelizumab/ocrelizumab placebo but were not randomized
- Patients who received ocrelizumab/placebo other than the group to which they were randomized at any point during the study
- Patients who received study drug that had been mishandled (e.g., incorrect storage temperature) and was not approved for subsequent use

4.1.4 Safety Population

The safety population will include all patients who received at least one dose of study drug (ocrelizumab or placebo). Randomized patients who receive the incorrect therapy

will be summarized in the group according to the therapy actually received, as listed below.

- Patients randomized to placebo who receive any ocrelizumab treatment and patients randomized to ocrelizumab who receive both placebo and ocrelizumab will be assigned to the ocrelizumab group.
- Patients randomized to ocrelizumab who receive placebo only will be assigned to the placebo group.
- Patients who are not randomized but who receive study drug will be included in the safety population and will be summarized as follows:
  - Patients who receive any ocrelizumab treatment will be assigned to the ocrelizumab group.
  - Patients who receive only placebo will be assigned to the placebo group.

4.1.5 OLE Safety Population

The OLE safety population will include all patients who received at least one dose of OLE study drug. Patients randomized to placebo who received at least one dose of ocrelizumab will be summarized in the ocrelizumab group consistent with the safety population.

4.2 OVERVIEW OF ANALYSES PLANNED FOR THE EXTENSION PERIODS

Overview of outcome measures and analyses performed for the extension periods in addition to the analyses performed for the primary CSR:

Period (Population) / Analyses	Extended controlled (ITT)	OLE (OLE ITT)	Combined extension (ITT)
Study conduct Patient Disposition (All patients)			X
Treatment group comparability Summary of demographics		X	
Baseline Disease Characteristics & baseline MRI Assessments		X	
Exposure to study drug Exposure to Ocrelizumab/Placebo			X
Confirmed disability progression /EDSS/T25-FW/9-HPT Time to Onset of Confirmed Disability Progression for at Least 12 weeks	X	X	X
Time to Onset of Confirmed Disability Progression for at Least 24 weeks	X	X	X

Period (Population) / Analyses	Extended controlled (ITT)	OLE (OLE ITT)	Combined extension (ITT)
Time to composite confirmed disability progression (EDSS or T25-FW or 9-HPT) for at Least 12 weeks	X	X	X
Time to composite confirmed disability progression (EDSS or T25-FW or 9-HPT) for at Least 24 weeks	X	X	X
Time to Onset of Confirmed Disability Progression Confirmed for ≥48 weeks	X	X	X
Time to composite confirmed disability progression for at least 48 weeks	X	X	X
Time to Onset of Confirmed EDSS ≥7 for at least 24 weeks	X	X	X
Time to Onset of Confirmed EDSS ≥7 for at least 48 weeks	X	X	X
Time to confirmed 20% increase in T25-FW for at least 12 weeks	X	X	X
Time to confirmed 20% increase in T25-FW for at least 24 weeks	X	X	X
Time to confirmed 20% increase in T25-FW for at least 48 weeks	X	X	X
Time to confirmed 20% increase in 9-Hole Peg for at least 12 weeks	X	X	X
Time to confirmed 20% increase in 9-Hole Peg for at least 24 weeks	X	X	X
Time to confirmed 20% increase in 9-Hole Peg for at least 48 weeks	X	X	X
Change in T25-FW speed from baseline (MMRM)	X	X	X
Change in EDSS from baseline (MMRM)	X	X	X
Change in MSFCS from baseline (MMRM)	X	X	X
MRI measures		X	
Ratio and percent change in T2 lesion volume from baseline		X	X
Ratio and percent change in non-enhancing T1 lesion volume from baseline		X	X
Number of Gadolinium-enhancing T1 lesions by visit			X
Number of new and enhancing T2 hyperintense lesions by visit			X
Percent change in total brain volume from Baseline		X	X
Percent change in total brain volume from Week 24			X
Percent change in cortical gray matter volume from Baseline		X	X

Period (Population) / Analyses	Extended controlled (ITT)	OLE (OLE ITT)	Combined extension (ITT)
Percent change in cortical gray matter volume from Week 24			X
Percent change in white matter volume from Baseline		X	X
Percent change in white matter volume from Week 24			X
Annualized Percentage Change in T1 lesion volume from Baseline to OLE end during the Extended Controlled Treatment and OLE Period		X	X
Annualized Percentage Change in T2 lesion volume from Baseline to OLE end during the Extended Controlled Treatment and OLE Period		X	X
Annualized Percentage Change in Total Brain Volume from Baseline to OLE end during the Extended Controlled Treatment and OLE Period		X	X
Annualized Percentage Change in Cortical Grey Matter Volume from Baseline to OLE end during the Extended Controlled Treatment and OLE Period		X	X
Annualized Percentage Change in White Matter Volume from Baseline to OLE end during the Extended Controlled Treatment and OLE Period		X	X
Descriptive statistics: SF-36, MFIS, NfL			X
Safety Summary: Additional safety data collected during ECP and OLE	X	X	X

9-HPT = 9-hole peg test; ECP = extended controlled treatment period; EDSS = Expanded Disability Status Scale; ITT = intent-to-treat; MFIS = Modified Fatigue Impact Scale; MMRM = Mixed-Effect Model Repeated Measures; MRI = magnetic resonance imaging; MSFCS = Multiple Sclerosis Functional Composite Scale; NfL = Neurofilament Light chain; OLE = open-label extension;T25-FW = timed 25-foot walk.

All analyses performed for the primary CSR are included in this statistical analyses plan. In general, the analyses performed on the extension periods follow the same analyses rules as for the primary CSR. Exceptions have been added in the respective sections.

4.3 ANALYSIS OF STUDY CONDUCT

The following analyses will be conducted to evaluate the study conduct:

- Summary of protocol deviations
- Summaries of ITT, per protocol defined, and safety populations, including numbers of patients in each population, and reasons for exclusion from the per-protocol population

- Summary of subject disposition, including the number of treatment doses received, and the number of patients entering into SFU
- Summary and Kaplan-Meier plots of:
  - Time to discontinuation of study treatment during the blinded treatment period. (All other patients will be censored at their last assessment during the blinded treatment period)
  - Time to discontinuation from the study (All other patients will be censored at their last assessment during the study)

Extension Periods

The summary of subject disposition will be updated using additional data from the ECP and OLE periods, including the number of patients entering into and withdrawing from OLE and into/from OLE SFU.

4.4 ANALYSIS OF TREATMENT GROUP COMPARABILITY

For continuous variables, the mean, median, standard deviation (SD), and minimum and maximum values will be calculated. For categorical variables, the number and percentage in each category will be displayed. For each item in the following lists, the units and categories to be used are indicated in parentheses and separated by commas. All durations are calculated with respect to the date of randomization, if not stated otherwise.

Demography and stratification factors based on electronic Case Report Form (eCRF) data:

- Age (years) at baseline
- Age stratification category ($\leq$45, $>$45 years)
- Sex (male patients, female patients)
- Race (White, Black or African American, Other)
- Ethnicity (Hispanic or Latino, nonHispanic or Latino)
- Weight (kg)
- Body mass index (measured in kg/m^2)
- Region stratification category (U.S. vs. ROW)

A summary of discrepancies in the age and region values (stratification factors) between the IxRS and the eCRF will be provided.

Baseline disease characteristics:

- Baseline EDSS (continuous)
- Baseline score in FS for each category (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, cerebral [or mental]; rated 0−5 or 0−6 depending on the domain of FS) and for ambulation (categorical; rated 0−12)

- Baseline Multiple Sclerosis Functional Composite (MSFC) score (continuous; raw results for each component)

MS disease history:

- Duration since MS symptom onset (calculated in years; i.e., divide by 365.25)
- Duration since MS symptom onset category (≤ 3 years, 3 to ≤ 5 years, 5 to ≤ 10 years, > 10 years)
- Duration since onset of PPMS diagnosis (calculated in years up to randomization date)
- Treatment with any MS disease-modifying therapy prior to the baseline visit (yes [treatment experienced patients], no [naive patients])
  - MS disease-modifying therapy includes Interferons, Glatiramer acetate, Natalizumab, Fingolimod, Dimethyl fumarate, Teriflunomide, Alemtuzumab, Mitoxantrone, Rituximab, Normal Immunoglobulin, Mycophenolate mofetil, Azathioprine, Immunotherapies for MS
  - Patients who received steroids for MS prior to the baseline visit [yes, no]

Baseline MRI data:

- Number of gadolinium-enhancing T1 lesions at baseline (continuous and categorical analysis [0, 1, 2, 3, ≥ 4])
- Whole-brain volume at baseline
- T2 volume at baseline
- Number of T2 lesions at baseline (continuous and categorical analyses [0-5, 6-9, > 9])

Extension Periods

- For OLE, demography and stratification factors, baseline disease characteristics, MS disease history and baseline MRI data will be summarized using the information collected at OLE baseline as described in the previous section, but subsetting for the OLE ITT population.
- OLE baseline score (EDSS, MRI, etc.) is the latest score collected prior to (including) OLE 1st Dose and after (including) Week 120. If the patient had no measurements within the defined time window, then set the OLE baseline score to be missing.

4.5 EFFICACY ANALYSIS

All statistical hypotheses for the primary and secondary endpoints and treatment comparisons will be tested at the 5% significance level ($\alpha = 0.05$) against two sided alternatives.

Apart from EDSS, for all other assessments, the baseline value will be used as the last non-missing value on or before the date of the first infusion of study drug. Baseline

EDSS is calculated differently. The derivation of baseline EDSS and time to onset of CDP is discussed in Section 4.5.1.

All primary and secondary efficacy endpoints will be stratified by region and age for analysis. The regional stratifications are U.S. and ROW. The age categories are $\leq$45 years and $>$45 years.

Extensions Periods

The analyses of the ECP and the combined ECP+OLE periods will be using the same baseline value as the primary analysis. For the analyses of the OLE period, the OLE baseline score is used.

4.5.1 Primary Efficacy Endpoint

Time to onset of CDP for at least 12 weeks (12-week CDP) during the double-blind treatment period

Significance Level

The null hypothesis will be tested at the $\alpha = 0.05$ level (two sided test).

The hypotheses to be tested are:

- H0 (null hypothesis): There is no difference in the time to onset of CDP between the ocrelizumab and placebo groups.
- H1 (alternative hypothesis): There is a difference in the time to onset of CDP between the ocrelizumab and placebo groups.

Definition

The time to onset of CDP (12 week confirmation [days]) is defined as the time from baseline to the onset of the first disability progression that is confirmed at the next regularly scheduled visit $\geq$12 weeks after the initial disability progression. If the patient has an infection, dosing may not occur on the Day 1 visit. Baseline for the time to onset of confirmed disability is the date of randomization, independent of the date of first dosing. For example, a subject with delayed dosing may receive the first dose 2 weeks after the baseline visit; if an EDSS score is recorded between the randomization date and the date of the first dose, this value will be considered for the date of initial disability progression. Disability progression is defined as an increase of $\geq$1.0 point from baseline EDSS score if the baseline EDSS value is $\leq$5.5 points (inclusive) or an increase of $\geq$0.5 points if the baseline EDSS value is $>$5.5 points. Assessments within 30 days after a protocol defined relapse will not be used for confirmation of CDP. The non-confirmatory EDSS assessments (between the initial disability progression and the confirmation of disability progression should also fulfill the requirements of the progression (see Section 2.1.1). Otherwise, the initial disability progression is not confirmed.

The baseline EDSS value is the average score of the EDSS assessment at screening and baseline (Day 1 visit) up to and including the date of randomization. If one of the values is missing, the non-missing values will be used as baseline.

An independent examining investigator at each study site will assess EDSS for all patients at the site at screening, baseline, every 12 weeks (regularly scheduled visit) during the blinded treatment period of the study, during the SFU period, at any unscheduled visits, and at withdrawal from treatment and end of study visits. Additional EDSS assessments for individual patients may be requested between visits (i.e., during an MS relapse).

The examining investigator is not the physician responsible for the patient care (the treating investigator).

The EDSS is based on a standard neurological examination; the seven categories of the EDSS representing FS (pyramidal, cerebellar, brainstem, sensory, bowel and bladder, visual, and cerebral [or mental], plus "other") are rated and scored (collectively, FSs), in addition to an ambulation score (0−12). Each domain of the FS is an ordinal clinical rating scale from 0 to 5 or 6. These ratings are then used in conjunction with observations and information regarding ambulation and the use of assistive devices (which will also be scored) to determine the EDSS score. The EDSS is a disability scale that ranges in 0.5point steps from 0 (normal) to 10.0 (death).

Analysis Methods

The ITT population analysis will be presented.

Patients who did not have initial disability progression at the time of primary database lock, time of early discontinuation, or loss to follow-up will be censored at the date of their last EDSS assessment that occurred during the treatment period.

There are various options to obtain EDSS assessments (confirmatory and non-confirmatory) and to record how the discontinuation from treatment and study can influence the confirmation of the disability progression. The following rules apply for all patients with initial disability progression:

- Initial disability progression must occur when the patient is still on treatment.
- All non-confirmatory EDSS assessments (if any) after initial disability progression and up to and including the confirmatory EDSS assessment should also fulfill the requirements of the progression as defined above. Otherwise an initial disability progression event cannot be confirmed.

Patients who experience initial disability progression will be treated as described in Table 2).

Table 2 Censoring Algorithm of Patients after Initial Disability Progression

Availability of EDSS Assessment for Patient with Initial Disability Progression	Recorded Disability Progression Status
Scenario A: Confirmatory EDSS assessment on treatment or after withdrawal from treatment	Confirmed disability progression
Scenario B: No confirmatory EDSS assessment; patient on treatment at the time of primary database lock	Censored at last EDSS assessment
Scenario C: Discontinues from treatment and is lost to follow-up; i.e., no available EDSS assessment at ≥84 days after initial disability progression	Confirmed disability progression

EDSS=Expanded Disability Status Scale.

The time to onset of CDP for the ocrelizumab arm and the placebo arm will be compared using a two-sided log-rank test stratified by geographic region (U.S. vs. ROW) and age (≤45 vs. >45 years). The proportion of patients with CDP at predefined timepoints (i.e., 48 weeks, 96 weeks, etc.) will be estimated using Kaplan-Meier methodology. The overall hazard ratio will be estimated using a stratified Cox regression model with the same stratification factors used in the aforementioned stratified log-rank test.

Sample SAS code can be found below (SAS code is regarded as "draft" until fully validated at the analysis stage):

```
* Log-rank test;
    proc lifetest data=CDP12;
        time CDP12TIME*CNSR(1);
        strata REGION AGE45 / group=ARMCD;
        ods output HomTests =_logrank;
    run;

* Cox regression model;
proc phreg data=CDP12;
        class ARMCD / descending;
        model CDP12TIME*CNSR(1) = ARMCD / rl;
        strata REGION AGE45;
        ods output ParameterEstimates=est;
    run;
```

Extension periods

This analysis will be repeated to include the EDSS assessments performed during the ECP and OLE periods for ECP and the combined double blind (DB)+OLE periods, with the same baseline as for the analysis during DB. Analysis containing only data from OLE will be performed considering the OLE baseline.

Patients discontinuing from either treatment periods (ECP or OLE) after initial disability progression and without confirmatory EDSS assessment will be imputed.

For the OLE period, initial disability progression has to occur while the patient is on OLE treatment. Confirmation of progression could occur at scheduled OLE or SFU visits after the OLE.

The same principle applies to all time to CDP endpoints below.

4.5.2 Secondary Efficacy Endpoints

The following secondary efficacy endpoints will be tested in the same hierarchical order listed here, if the primary endpoint and each preceding endpoint have reached the significance level of 0.05.

4.5.2.1 Time to Onset of Confirmed Disability Progression Confirmed for ≥24 Weeks

In addition to the primary analysis, the time to onset of CDP will be analyzed in the ocrelizumab and placebo treatment arms with use of a 24 week confirmation window for disability progression. The analysis of EDSS progression will be conducted as described for the primary analysis, with the exception that the time to onset of CDP (24 week confirmation) is defined as the time from baseline to the first disability progression that is confirmed at the next regularly scheduled visit ≥24 weeks (≥161 days) after the initial disability progression.

Extension Periods

This analysis will be repeated to include the EDSS assessments performed during the ECP and OLE periods for ECP and the combined DB+OLE periods, with the same baseline as for the analysis during DB. Analysis containing only data from OLE will be performed considering the OLE baseline.

4.5.2.2 Change in 25-Foot Timed Walk from Baseline to Week 120

The 25-foot timed walk is one of the three primary measures in an Multiple Sclerosis Functional Composite Scale (MSFCS) score and is the mean time from the two trials performed on the same visit. Detailed description on the derivation of the MSFCS score can be found in the MSFC Administration and Scoring Manual (Fischer 2002).

If the test results are not available due to a “physical limitation,” the maximum possible value for the scale (180 seconds) will be imputed. If one of the two trials is not available and not missing due to a “physical limitation,” the result from the other trial will be used to impute the missing value.

For outliers, the following rules will be applied. Values outside the lower and upper bounds will be treated as missing, and the imputation rule will be applied as defined above. The lower bound is 2.2 seconds (Bohannon 1997), and the upper bound is 180 seconds, according to the MSFCS Manual.

For the assessment of differences in the change in the 25-foot timed walk from baseline up to Week 120, a nonparametric ranked analysis of covariance (ANCOVA) will be

performed, with the ranked percent change in the 25-foot timed walk as the outcome variable and the ranked baseline value as the covariate, adjusting for geographical region (U.S. vs. ROW) and age ($\leq$45 vs. $>$45 years).

The last-observed value available before treatment discontinuation will be carried forward to impute the missing values. The baseline value will be carried forward for patients who did not have evaluable post-baseline observations. Patients with missing baseline value will be excluded from the analysis.

In order to obtain estimates of the treatment effect, a Mixed-Effect Model Repeated Measures (MMRM) analysis will be used. This model does not use the last observation carried forward (LOCF) method. The MMRM analysis incorporates post-randomization data collected at scheduled visits up to 120 weeks of treatment, and it will be used to assess all data collected over time, with consideration of the variance−covariance matrix of the repeated measures. This method allows for inclusion of data from patients with incomplete data from some scheduled timepoints.

The model will be implemented in SAS using PROC MIXED and will include the log transformed ratio of post-baseline timepoints to baseline in the 25-foot timed walk at each visit as the dependent variable. The fixed effects in the model will include independent variables of randomized treatment, visit (nominal post baseline visits as per the Schedule of Assessments), baseline-by-visit interaction and treatment-by-visit interaction, along with the following baseline covariates, log-transformed 25-foot timed walk at the baseline visit, geographical region (U.S. vs. ROW), and age ($\leq$45 vs. $>$45). Visit will be treated as a repeated variable within a patient. Patient, treatment and visit will be treated as factor variables. An unstructured variance−covariance structure will be applied to model the within-patient errors. The Restricted Maximum Likelihood (REML) method will be used for estimates of variance components. Denominator degrees of freedom will be estimated using Satterthwaite's approximation.

To estimate the difference between the ocrelizumab and placebo groups in change from baseline to Week 120, a treatment-by-visit interaction contrast will be constructed (i.e., the treatment group contrast at Week 120). On the basis of this analysis, back transformed estimates of least square means and the 95% confidence interval (CI) for the ratio of 25-foot timed walk at each visit relative to the baseline will be reported.

Graphical presentations for least square means and 95% CIs will be used to illustrate trends over time.

Sample SAS code can be found below (SAS code is regarded as “draft” until fully validated at the analysis stage):

```
*ranked ANCOVA;
    proc sort data=MSFC_locf out=ancova;
        by REGION AGE45;
    run;

    proc rank data=ancova out=ranks;
        by REGION AGE45 ;
        var lbfwt lcfwt;
    run;

    proc reg data=ranks;
        by REGION AGE45;
        model lcfwt=lbfwt ;
        output out=residual r=reside;
        quit;

    proc freq data=residual noprint ;
        tables region*AGE45*ARMCD*reside/cmh2 scores=modridit;
    run;

*MMRM;
    proc mixed data=MSFC method=REML;
        class REGION AGE45 ARMCD VISIT USUBJID;
        model log(AVG25F/BASE)  =  log(BASE)  REGION
        AGE45 ARMCD VISIT log(BASE)*VISIT  ARMCD*VISIT /
         ddfm= satterthwaite;
        repeated VISIT / type=un subject = USUBJID;
        lsmeans ARMCD*VISIT / pdiff cl;
        ods output lsmeans = lsm;  * contains the adjusted means;
        ods output diffs = dif;  * contains treatment differences;
    run;
```

Extension Periods

This analysis will be modified to include data collected during the ECP and OLE periods. Change from baseline during ECP, change from baseline until the last mature visit during the combined ECP+OLE periods and change from OLE baseline until the last mature visit during the OLE period will be analyzed. The last mature visit is defined as the visit where the majority of patients have completed prior to the data cutoff date.

The model will include the 1/x transformed value of the 25-foot timed walk at each visit as the dependent variable, instead of the log-transformation applied previously.

For the combined ECP and OLE periods, visits will be backdated in order to have continuous time visits post ECP. Below algorithm will be considered for backdated OLE visits.

Backdate Algorithm (continuous OLE visits):

- All scheduled OLE labeled visits to be windowed to the nearest 12-week visit following the scheduled visits during the ECP.

- For each OLE patient, create an artificial Schedule of Assessment (SoA): a visit every 12 weeks in continuation to the ECP.
- Assign a new visit label from the closest visit from artificial SoA to OLE visit.
- All OLE Week yy visits to be replaced by Week xx, in continuation to the visit labels during the ECP.
- Use the new visit labels (no OLE Week yy visit label) for the analysis purpose.
- If multiple OLE visits are windowed to the same 12 week visit, then the latest chronological visit will be used for the analyses.

4.5.2.3 Change in Total Volume of T2 Lesions on MRI Scans of the Brain from Baseline to Week 120

For the assessment of differences in the percent change in total volume of T2 lesions from baseline up to Week 120, ranked ANCOVA and MMRM analyses will be undertaken using the same approach as described above in Section 4.5.2.2, except that the baseline covariate for the MMRM will be baseline log-transformed T2 lesion volume, geographical region (U.S. vs. ROW), and age ($\leq$45 vs. $>$45 years).

T2 lesion volumes below the lower limit of quantification of 0.009 cm^3 (3 voxels) will be set to this limit.

Extension Periods

This analysis will be extended to include data collected during the OLE period. Ratio relative to baseline until the last mature visit during the combined ECP+OLE period and ratio relative to OLE baseline until the last mature visit during the OLE period will be analyzed. Note that no additional MRI assessments were performed between 120 weeks and OLE start, hence analysis for the ECP period is the same as that for the DB period.

4.5.2.4 Percent Change on MRI from Week 24 to Week 120 in Total Brain Volume

For the assessment of differences in the mean percentage change in brain volume on MRI scans from Week 24 to Week 120, an MMRM analysis will be undertaken using the same approach as described above in Section 4.5.2.2, except:

- The response variable will be the percentage change from Week 24 to subsequent visits (Weeks 48, and 120). Baseline brain volume data will not be a feature in this analysis.
- Covariates will be brain volume at Week 24, geographical region (U.S. vs. ROW), and age ($\leq$45 vs. $>$45 years).

The choice of Week 24 as “baseline” in this analysis, rather than the baseline visit, is a widely accepted approach to establish a stable baseline and avoid the physiological effect of pseudo-atrophy.

Extension Periods

This analysis will be extended to include data collected during the OLE period. Percentage change from Week 24, baseline and OLE baseline until the last mature visit will be analyzed.

4.5.2.5 Change from Baseline in Quality of Life as Measured by the SF36 (Physical Component)

The SF-36v2 is a 36-item, self- reported, generic measure of quality of life that has been widely used in multiple disease areas. It is composed of eight health domains: Physical Functioning (PF), Role-Physical (RP), Bodily Pain (BP), General Health (GH), Vitality (VT), Social Functioning (SF), Role-Emotional (RE), and Mental Health (ME). Based on these domain scores, the Physical Component Summary (PCS) Score and the Mental Component Summary (MCS) Score will be computed. The standard form of the instrument, with a 4-week recall, is being administered in the studies.

Scoring and rules for missing items will follow the SF-36v2 User's Manual. In brief, scoring for each health domain scale involves (a) recoding item response values, (b) summing recoded response values for all items in a given scale to obtain the scale raw score, and (c) transforming the scale raw score to a 0−100 score. The PCS score is computed by (a) multiplying each health domain z score by a scale-specific physical factor score coefficient, (b) summing the resulting products, and (c) converting the product total to a T score. The MCS score is computed in the same manner, instead using scale-specific mental factor score coefficients.

For the assessment of differences in the mean change in PCS score from baseline up to Week 120, MMRM analyses will be undertaken using the same approach as described above in Section 4.5.2.2, except baseline covariates here will be as follows: baseline PCS score, geographical region (U.S. vs. ROW) and age (≤45 years, >45 years).

Extension Periods

Additional data from the OLE period will be descriptively analyzed. Note that no additional SF-36 assessments were performed between 120 weeks and OLE start.

4.5.3 Exploratory Efficacy Endpoints

The MS guideline of the European Medicines Agency (EMA, 26 March 2015) suggests a number of primary and secondary efficacy outcome measures. On the basis of these recommendations, the following exploratory efficacy outcomes will be analyzed:

- MRI derived parameters: number of gadolinium-enhancing T1 lesions and number of new or enlarging T2 lesions. For the combined ECP and OLE periods, these parameters will be analyzed by scheduled visit as follows:
  - Negative binomial model adjusted by Geographical Region (US vs. ROW) and Age (≤45, >45 years) for the baseline visit.

  - Negative binomial model adjusted by Baseline T1 Gd Lesion (present or not) or Baseline T2 lesion count, Geographical Region (US vs. ROW) and Age ($\leq$45, $>$45 years) for the post-baseline visits.
  - If the negative binomial model doesn't converge due to small number of lesion counts, unadjusted rate and exact poisson CI will be presented.

- Cognitive impairment: change from baseline to Week 120 in the Paced Auditory Serial Addition Test (PASAT).
- Percent change from baseline to Week 120 in total brain volume, white matter volume and cortical gray matter brain volume and the percent change from Week 24 to Week 120 in white matter volume and grey matter brain volume. For all parameters, analyses will be extended to include data collected during the OLE period where percentage change from Week 24, baseline and OLE baseline until the last mature visit will be analyzed.

Additional exploratory outcomes that will be analyzed include the following:

- Proportion of patients with CDP at Week 120
- Functional impairment: change from baseline in EDSS score, and change from baseline in MSFCS. For both parameters, analyses will be extended to include data collected during the ECP and OLE periods. Change from baseline during ECP, change from baseline until the last mature visit during the combined ECP+OLE periods and change from OLE baseline until the last mature visit during the OLE period will be analyzed.
- The change in fatigue, as measured by the Modified Fatigue Impact Scale (MFIS) total score and subscale scores (Physical impact, Cognitive impact, and Psychological impact) from baseline to Week 120. Additional data from the OLE period will be descriptively analyzed.
- The change in quality of life, as measured by the SF-36v2 MCS score from baseline to Week 120. Additional data from the OLE period will be descriptively analyzed.
- Time to composite confirmed disability progression (cCDP) over the treatment period, defined as an increase in EDSS that is sustained for at least 12 weeks (0.5 or 1, same criteria as for the primary endpoint time to 12-week CDP) or a 20 percent increase in 25-foot timed walk that is sustained for at least 12 weeks or a 20 percent increase in 9-hole peg test that is sustained for at least 12 weeks. This analysis will be repeated to include the EDSS assessments performed during the ECP and OLE periods for ECP and the combined DB+OLE periods, with the same baseline as for the analysis during DB. Analysis containing only data from OLE will be performed considering the OLE baseline.

Additional endpoints for the extension periods:

- Change from baseline in non-enhancing T1 lesion volume will be analyzed similarly to the T2 lesion volume. Ratio relative to baseline until the last mature visit during the combined ECP+OLE period and ratio relative to OLE baseline until the last mature visit during the OLE period will be analyzed.

- Time to Onset of CDP confirmed for ≥48 weeks

Time to onset of CDP will be analyzed with use of a 48 week confirmation window for disability progression. The analysis will be conducted as described for the primary analysis, with the exception that the time to onset of CDP (48 week confirmation) is defined as the time from baseline to the first disability progression that is confirmed at the next regularly scheduled visit ≥48 weeks (≥329 days) after the initial disability progression.

The analysis will be performed for ECP and the combined DB+OLE periods, with the same baseline as for the analysis during DB. Analysis containing only data from OLE will be performed considering the OLE baseline.

- Time to Onset of cCDP for at least 24 weeks or 48 weeks

Time to onset of cCDP will be analyzed with use of a 24 or 48 week confirmation window for disability progression. The analysis will be conducted as described for the primary analysis.

The analyses will be performed for ECP and the combined DB+OLE periods, with the same baseline as for the analysis during DB. Analysis containing only data from OLE will be performed considering the OLE baseline.

- Time to Onset of Confirmed EDSS ≥7 for at least 24 weeks or 48 weeks

Time to onset of confirmed EDSS ≥7 will be analyzed with EDSS ≥7 (requiring wheelchair) to be defined as a disability progression event that is sustained for at least 24 or 48 weeks. The analysis will be conducted as described for the primary analysis.

The analyses will be performed for ECP and the combined DB+OLE periods, with the same baseline as for the analysis during DB. Analysis containing only data from OLE will be performed considering the OLE baseline.

- Annualized Percentage Change in MRI assessments from Baseline during the combined ECP and OLE Periods

Due to the event driven study design and no additional MRI assessments were performed between 120 weeks and OLE start, the time gap between 120 weeks and OLE start varies from patient to patient. Hence, an annualized change in all MRI assessments, including the total brain volume, white matter volume, cortical gray matter volume, T1 lesion volume and T2 lesion volume is analyzed for the combined ECP and OLE periods, which takes into account the variable time difference between visits for each patient.

At each scheduled visit during the ECP and OLE periods, calculate the annualized change from the previous visit as follows:

$$CHG = \frac{\frac{AVAL - lag(AVAL)}{lag(AVAL)} \times 100}{\left(ADT - lag(ADT)\right)/365.25}$$

where AVAL corresponds to the MRI value at the current visit, and ADT corresponds to the date of the current visit.

For T1 and T2 lesion volume, the log transformation is applied, and 0 values are imputed to 0.009

$$LOGCHG = \frac{\log(AVAL) - \log(lag(AVAL))}{\left(ADT - lag(ADT)\right)/365.25}$$

The OLE MRI assessments are backdated as follows:

- Create an SoA from Week 120 with a visit every 48 weeks, e.g., Week 168, Week 216, etc
- All OLE visits prior to Week 168 will be assigned to Week 168 ≥New visit label
- For all remaining OLE visits, they will be assigned to the nearest Week xxx with a time window of ±24 weeks ≥New visit label
- Use the new visit labels (no OLE Week yy visit label) for the analysis
- If multiple OLE visits are windowed to the same Week xxx visit, then the latest chronological visit will be used for the analyses

- Neurofilament Light chain (NfL) data collected during the ECP and OLE periods will be summarized descriptively. Summary statistics of the absolute NfL values will be presented at all scheduled visits along with the percentage change from baseline. A corresponding plot will be produced. The analyses will be performed for the ECP period only as well as for the combined ECP and OLE periods.

4.5.4 Sensitivity Analyses

The following sensitivity analyses of the primary endpoint will be conducted according to the analysis described in Section 4.5.1:

1. The primary analysis will be repeated using the per-protocol population as the analysis population. The below sensitivity analyses will be done with the ITT population.
2. A sensitivity analysis using multiple imputations will be performed for the ITT population to explore the potential influence of informative censoring on the results of the primary efficacy analyses. Censoring for a reason that is not independent of that patient's prognosis is called "informative censoring." The influence of

informative censoring needs to be explored, because for time-to-event endpoints, the ITT principle requires that essentially all randomly assigned patients be observed up to the endpoint or up to the end of the study. When a patient's follow-up is censored X number of months after random assignment, then, in the computation of the Kaplan-Meier estimates, log-rank or Cox regressions, that patient's outcome after X number of months is assumed to have the same outcome as the other patients in their treatment group who also are free of the outcome at X number of months and who remain under follow-up beyond X number of months. Thus, unless the reason for being censored is independent of that patient's prognosis, failure to observe that patient until occurrence of his study endpoints could lead to significant bias as well as increased variability in the evaluation of treatment effects. To censor time to CDP at dropout due to premature withdrawal and lost to follow-up after an initial disability progression is potentially to favor the most toxic or less efficacious treatment and, consequently, should be avoided whenever possible. Thus the purpose of this sensitivity analysis is to determine whether informative censoring has any influence on the primary endpoint.

Multiple imputations will be used to impute the events for patients who had initial disability progression and then discontinued treatment with no confirmatory EDSS assessments. Multiple imputation inference involves three distinct phases:

- The missing data are filled in m times to generate m complete data sets. Instead of filling in a single value for each missing value, multiple imputation replaces each missing value with a set of m plausible values that represent the uncertainty about the right value to impute.
- The m complete data sets are analyzed using standard statistical analyses.
- The results from the m complete data sets are combined to produce inferential results.

Fifty percent of the patients who had initial disability progression and then discontinued the treatment with no confirmatory EDSS assessments will be randomly assigned to have an event at the time of their initial disability progression; the other 50% of these patients will be censored at their last EDSS assessment, under the assumption that 50% of the patients with initial disability progression will later be confirmed at 12 weeks. A total of 1000 ($m=1000$) imputed datasets will be produced. The primary model described in Section 4.5.1 will be applied to each of the imputed datasets. Each imputed dataset will produce an estimate of the difference between ocrelizumab and placebo. The multiple imputation estimator of the difference between ocrelizumab and placebo is the average of the individual 1000 estimators. The variance of the estimator is the combination of the between- and within-imputation variability (Carpenter and Kenward 2007).

The 50% confirmation rate for patients with initial disability progression was estimated on the basis of previous studies conducted in patients with PPMS (Weinshenker et al. 1996; Olympus, Study U2786g, data available upon request).

3. A sensitivity analysis, in which the patients who had initial disability progression and then discontinued the treatment with no confirmatory EDSS assessments will be considered to not have CDP, will be performed for the ITT population. This sensitivity analysis assumes that all patients lost to follow up after initial disability progression events will not have reached CDP. Hence, it will give an estimate of the maximal effect of informative censoring on the parameter estimates.
4. The influence of early progression events on treatment effect will also be evaluated by omitting the EDSS assessments performed between randomization and the Week 12 visit (≤83 days after randomization) with use of the ITT population.
5. The study originally planned to enroll approximately 630 patients who were to be followed for at least 120 weeks. The study actually enrolled 732 patients (last patient enrolled on 27 December 2012). To explore the effect of the higher enrollment number, the analysis of the primary efficacy endpoint will be performed using the first 630 patients on the basis of their randomization dates (ITT population) at the time of the primary analysis.
6. The primary analysis described above will be adjusted with additional strata for the baseline presence of gadolinium-enhancing T1 lesions (present or absent) and baseline EDSS (≤5.5 vs. >5.5).
7. PPMS is characterized by a progressive course from disease onset typically without superimposed discrete clinical attacks or relapses (Ebers 2004). The primary analysis on CDP for ≥12 weeks, described above will be performed excluding patients who have any clinical relapses during the study. Clinical relapses include protocol-defined relapses.
8. For the time to onset of CDP for ≥24 weeks, the influence of early progression events on treatment effect will also be evaluated by omitting the EDSS assessments performed between randomization and the Week 12 visit (≤83 days after randomization) with use of the ITT population.
9. For the time to onset of CDP for ≥24 weeks, the analysis will be performed excluding patients who have any clinical relapses during the blinded treatment period.

4.5.5 Subgroup Analyses

The primary and key secondary endpoints (time to onset of 12- and 24 week CDP, change in 25-foot timed walk from baseline to Week 120, and change in total volume of T2 lesions, respectively) will be summarized by the following subgroups:

- Age (≤45 vs. >45 years)
- Sex (male patients vs. female patients)
- Baseline EDSS (≤5.5 vs. >5.5)
- Region (U.S. vs. ROW)
- Presence or absence of gadolinium-enhancing lesions at baseline MRI scan

- Prior MS disease-modifying therapies, with the exception of corticosteroids (yes vs. no)
- Duration since MS symptom onset (≤3 years, 3 to ≤5 years, 5 to ≤10 years, >10 years)
- Weight (≤75 vs. >75 kg at baseline)
- Body mass index (<25 vs. ≥25 kg/m^2, at baseline)

4.6 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

4.6.1 Pharmacokinetic Analyses

Nonlinear mixed effects modeling will be used to analyze the sparse sampling dose–concentration–time data of ocrelizumab. Patients who have measurable concentrations of ocrelizumab will be included in the PK analysis unless major protocol deviations or the unavailability of information (e.g., exact blood sampling time) occur that may interfere with PK evaluation. The PK data from this study may be pooled with other studies. Population PK parameters (clearances and volumes) will be estimated, and the influence of covariates such as age, sex, weight, ADAs, and baseline $CD19^+$ lymphocytes on these parameters will be investigated.

The model developed based on relapsing multiple sclerosis (RMS) data will be used to predict PK profiles in the PPMS study. If predictions match observations, this population PK model based on RMS data will be deemed acceptable to describe the PK in PPMS patients too. If predictions do not match observations, a population PK model based on the PPMS or pooled RMS+PPMS data will be developed.

Details of the mixed effects modeling analyses will be described in a Modeling and Simulation Analysis Plan, and results will be reported separately.

4.6.2 Pharmacodynamic Analyses

If possible, nonlinear mixed effects modeling will be used to investigate the relationship between ocrelizumab exposure and selected safety and efficacy parameters.

Exploratory analyses may be performed to assess the possible relationship between PD markers like B cells, pharmacokinetics, and clinical response; this will be described in a Modeling and Simulation Analysis Plan, and results will be reported separately.

4.7 SAFETY ANALYSES

The safety population is defined in Section 4.1.3. All safety parameters will be summarized and presented in tables on the basis of this safety population. Randomized patients who receive incorrect therapy will be summarized in the group according to the rules laid out in Section 4.1.3. Placebo patients who received any ocrelizumab in error will be assigned to the ocrelizumab group.

Patients who are not randomized but who receive study drug will be included in the safety population and summarized according to the therapy actually received.

This SAP describes the analyses of all available safety data collected until the end of the trial. The safety data will be summarized descriptively for the DBP, ECP, OLE and SFU periods.

4.7.1 Exposure of Study Drug

The amount of ocrelizumab or placebo infused at each infusion will be listed and summarized using descriptive statistics.

Definition of dose: A dose of ocrelizumab is given as two infusions administered 2 weeks apart.

Patients will be considered to have received a dose of treatment if at least part of one infusion of that dose (either Day 1 or Day 15 for dual infusions) was given. If a dose is completely missed instead of delayed, the next dose number will be based on the number of previous doses received.

The duration of observation for a patient will be calculated as follows:

(Date of last contact*−Date of first infusion in the first dose)+1

*Earliest 1) date of CCOD for the primary analysis reporting; 2) date of subject completed or discontinued early from the study completion end of study page; or 3) date of death.

The duration of observation, within a dose, is defined in a similar manner as follows:

(Day prior to first infusion in the n+1th dose*−Date of first infusion in the n^{th} dose)+1

*With the exception of the last dose received by the patient where the date of last contact is used as defined above. If the last contact is after the date of the clinical cutoff, the date of last contact will be the CCOD.

4.7.2 Adverse Events

Adverse events (AEs) will be defined as all AEs including infusion-related reactions (IRRs) and serious MS relapses, but excluding non-serious MS relapses. Therefore, those AEs recorded on the AE and IRR CRF pages will be included. Refer to Section 4.7.2.5 for further guidance on MS relapses considered SAEs.

For each recorded AE, the term entered by the investigator describing the event (the “reported term”) will be assigned a standardized term (the “Preferred Term” [PT]) and assigned to a superclass term on the basis of the MedDRA World Health Organization

(WHO) dictionary of terms. All analyses of AE data will be performed using the PTs unless otherwise specified.

For all summary tables, the AEs will be sorted by System Organ Class (SOC; in decreasing order of overall incidence) and then by PT (in decreasing order of overall incidence). All summaries and listings of AEs will be based on the safety population, unless stated otherwise. Summaries of AEs will be generated to summarize the incidence of treatment emergent AEs only. Treatment Emergent events are defined as those AEs with an observed or imputed date of onset on or after the start date of trial treatment. If the onset date of the AE is prior to the day of first dose, the AE will be considered treatment emergent only if the most extreme intensity is greater than the initial intensity (i.e., the intensity for a given AE increases and its end date is on or after the date of the first dose). An AE with a completely missing, nonimputed start date will be assumed to be treatment emergent unless the AE has a complete, nonimputed end date that is prior to the date of the first dose.

AEs will be assigned to a dose if the AE onset date is on or after the date of the first infusion of that dose but before the first infusion of the next treatment dose. AEs that are reported during the SFU period (including the B-cell monitoring period) will be assigned to the last dose received. Hence, the last dose for a patient may be of a variable length, from 24 to 72 weeks or longer. AEs that start prior to the first dose and worsen during treatment (i.e., treatment-emergent) will be assigned to the first treatment dose.

The number of patients who experienced a related AE will be summarized by SOC and PT. AEs will be summarized by SOC and PT by intensity grade. Multiple occurrences of the same event within a patient will be counted once at the greatest intensity/highest grade for this PT. For AEs leading to death, the most extreme intensity will be overwritten by Grade 5 (death). Any AEs and the SOC overall rows of the summary table will count patients according to AEs by intensity (grade).

All patient deaths, regardless of treatment received, will be listed.

SAEs will be defined as all SAEs including serious MS relapses and serious IRRs. The number of patients who experienced an SAE will be summarized by SOC and PT. Related SAEs will be summarized by SOC and PT. Additionally, the most frequent SAEs ($\geq$1%) will be presented by PT.

A patient may experience an AE that leads to the discontinuation of his/her study treatment. Discontinuation of study treatment for an AE may not necessarily lead to discontinuation from the study because the patient can enter the SFU period of the protocol. Only AEs that led to the discontinuation of study treatment are of interest. Patients who withdraw early from the study because of AEs will be summarized under disposition. The number of patients who experienced an AE that led to discontinuation of study treatment will be summarized by SOC and PT. The number of patients who

experienced an AE that led to modification or interruption of study drug will be summarized by SOC and PT.

For each treatment group, the incidence count for each AE PT will be defined as the number of patients reporting at least one treatment emergent occurrence of the event. The incidence rate will be calculated as the incidence count divided by the total number of patients in the population. Each table will also present the overall number of patients experiencing at least one AE and the total number of AEs reported (multiple occurrences of the same AE in 1 patient will be counted only once).

The rate per 100 patient-years by treatment group (along with the 95% CI) will be calculated for specific events of interest such as all AEs, SAEs, infections and serious infections, opportunistic infections (OIs), and malignancies (see Section 4.7.2.2, Section 4.7.2.3 and Section 4.7.2.4, respectively).

- The number of AEs per 100 patient-years is calculated as:

Total number of AEs / Total number of patient-years $\times$ 100

- The 95% CI for the number of AEs per 100 patient-years is calculated as follows (Sahai and Kurshid 1993):

Exact lower 95% confidence limit = chisq (p = 0.025, df = 2 Y) / (2T)

Exact upper 95% confidence limit = chisq [p = 0.975, df = 2(Y + 1)] / (2T)

where Y is the total number of AEs, T is total number of patient-years at risk and chisq (p,df) is the quantile of the upper tail probability of the X^2 distribution on df (degrees of freedom). This approach has the advantage of providing an estimate for the upper 95% confidence limit even when the total number of AEs is zero.

The rates per 100 patient-years will be summarized by treatment and dose and by treatment (including all doses).

4.7.2.1 Infusion-Related Reactions

An IRR and its corresponding symptoms are collected on the dedicated eCRF page.

The symptom(s) of an IRR and the IRR itself may be of different intensities. Because other symptoms can be recorded as free text on the eCRF page, symptoms will be coded in the MedDRA and summarized by PTs.

For IRRs, the number and percentage of patients with at least one infusion reaction will be presented per infusion (patients with multiple events within an infusion will count only once). In addition, the total number of IRRs will be summarized (multiple events will be counted). The total and percentage of events (based on the total number of patients

with at least one IRR) by most extreme intensity will be summarized per infusion and per dose. The number of serious IRRs will also be presented.

For multiple events in a given patient, the most extreme intensity will be used and the total number of events of each intensity will be equal to the total number of patients with at least one IRR if there are no missing extreme intensities. The total number of IRRs experienced by each subject will be summarized across the treatment doses.

IRRs are recorded on the CRF by the time of the event, (i.e., occurring during the infusion, after completion of the infusion while patient is in the clinic, or within 24 hours of completion of the infusion and the patient is not in the clinic). The number of patients with at least one IRR, the total number of IRRs, and the intensity of the IRR will be presented by the time of event.

Additionally, similar tables will be presented by the premedication subgroup (methylprednisolone alone, methylprednisolone plus analgesic/antipyretics, methylprednisolone plus antihistaminics, or methylprednisolone plus analgesic/antipyretic and antihistaminics).

Symptoms of IRRs and symptoms of serious IRRs will be presented by infusion. Symptoms of IRRs that led to modification of study drug or symptoms of IRR that led to discontinuation of study drug will also be presented. Symptoms of IRR experienced by patients during an infusion, after the infusion, and after the patient leaves the clinic will also be presented.

AEs other than IRRs experienced by the patient within 24 hours of an infusion will be presented by infusion. AEs will be included if the onset date is on the day of an infusion or the following day. The onset date of the AE will be matched to the date of an infusion; onset dates will not be imputed.

The concomitant treatments for IRRs will be summarized.

4.7.2.2 Infections

Infections will be defined from the AE data using the MedDRA SOC of “Infections and infestations.” During Primary CSR, AEs reported as an infection by the investigator will also be included. Infections will be classified according to the pathogen type (e.g., bacterial, fungal, viral, parasitical, unknown, or other).

During Primary CSR, an infection has been defined as serious if the event is a SAE or if the non-serious infection was treated with an IV anti-infective. For subsequent reporting, an infection has been defined as serious if the event is a SAE.

A listing will be presented of non-serious infections treated with an IV anti-infective during primary CSR. The number of patients who experienced an infection will be summarized by SOC and PT and by intensity.

The number of infections experienced by more than 5% of patients will be summarized by SOC and PT.

The number of patients who experienced a serious infection will be summarized by SOC and PT.

Identified pathogen codes will not be summarized but will only be listed within the listings of infections and serious infections. Infections and serious infections will be summarized by pathogen types.

The time to the first serious infection per patient will be summarized by treatment.

The rate per 100 patient-years by treatment group (along with the 95% CI) will be calculated for infections and serious infections overall and by dose based on the number of patient-years observation for the specific dose.

For infections and serious infections, the incidence rate ratio will be calculated. For these analyses, the incidence rate ratio with 95% CI based on the Poisson distribution exact method will be presented.

4.7.2.3 Opportunistic Infections

Opportunistic infections (OIs) will be defined using the Ocrelizumab specific MedDRA Term Selection (MTS) basket of “Opportunistic infections.”

The number of patients who experienced an OI will be summarized by SOC and PT.

The rate per 100 patient-years by treatment group (along with the 95% CI) will be calculated for OIs overall.

4.7.2.4 Malignancies

Malignancy and pre-malignancy AEs will be identified using the standard MedDRA query (SMQ) of “Malignant tumours (narrow)” and “Pre-malignant disorders,” respectively.

The number of patients with a malignancy will be summarized by SOC and PT. The pre-malignancy AEs will be listed.

The rate per 100 patient-years by treatment group (along with the 95% CI) will be calculated for malignancies.

4.7.2.5 Multiple Sclerosis Relapses

MS relapses will include all events recorded on the AE CRF pages that are classified as MS relapse.

Information related to a protocol defined relapse will be captured on a clinical relapse event CRF page. If the following criteria are satisfied, the clinical relapse will qualify as a protocol defined relapse:

- Check box labeled "Did symptoms persist for >24 hours and were not attributable..." on the clinical MS relapse event CRF page is checked.
- Increase in EDSS from a visit occurring at (or soon after) the date of onset of relapse of ≥ 0.5 compared with the previous EDSS; or for selected FS domains involved in the relapse event (pyramidal, ambulation, cerebellar, brainstem, sensory, or converted visual), a ≥ 2 point increase in one appropriate FS domain or a ≥ 1 point increase in two or more appropriate FS domains.
- No protocol defined relapse within 30 days before the date of onset of the clinical relapse.

For protocol defined relapses and clinically reported relapses (i.e., patients who experience any relapse), the number and percentage of patients experiencing an event will be summarized.

Clinical relapse will be considered an SAE when the relapse results in hospitalization for any reason other than routine treatment of the relapse (e.g., for a treatment course beyond the standard treatment described or when hospitalization is prolonged). These events will be listed and summarized.

4.7.2.6 Pregnancies

Pregnancy information will be summarized.

4.7.3 Magnetic Resonance Imaging Data

Non-MS pathology reported by the local safety radiologist on the CRF will be summarized by treatment group.

4.7.4 Laboratory Data

General Laboratory Evaluations

Abnormal laboratory outcomes will be reported. A summary of the number and percentage of patients with abnormal laboratory outcomes, along with each grade by laboratory parameter, will be summarized by treatment group for all laboratory assessments.

The absolute values and changes from baseline at each visit will be summarized for all laboratory assessments.

For the liver laboratory parameters, the number and percentage of patients with an elevated post-baseline AST or ALT level will be summarized by treatment.

CD19

The median CD19 cell count will be displayed graphically over time from the first infusion of study drug. Absolute CD19 counts and percent changes from baseline in the CD19 count will be summarized over time.

Immunoglobulins

The median immunoglobulin levels (IgA, IgG, IgM, and total Ig) will be displayed graphically over time from the first infusion of study drug. Absolute values, changes from baseline, and percent changes from baseline will be summarized over time. At each timepoint, the number and percentage of patients with immunoglobulin levels lower than the lower limit of normal will be presented.

Antibody Titers

If data are available, antibody titers for mumps, rubella, varicella, and *Streptococcus pneumoniae* will be summarized for the number and percentages of patients with a positive level by visit.

Anti-Drug Antibodies

ADAs, also called human anti-human antibodies, will be summarized over the blinded treatment period by treatment group. The baseline prevalence and post-baseline incidence of ADAs will be displayed. The number of patients with treatment-induced ADA will also be displayed. A table that summarizes ocrelizumab serum concentrations (µg/mL) at timepoints where ADA samples were collected and analyzed will be presented. A listing by treatment of antiocrelizumab antibody data will be presented for patients with at least one ADA sample.

HBV DNA

HBV DNA in patient core HBV antibody positive at screening will be listed.

4.7.5 Vital Signs

Vital sign and physical examination results, and ECG data will be included in individual patient listings.

Changes from study baseline in vital signs will be summarized by visit and group. Changes from pre-infusion baseline to post-infusion timepoints will also be summarized for each infusion.

4.7.6 Clinical Genotyping

The following genetic markers will be summarized descriptively. A summary of polymorphism ID, genotype frequency, and allele frequency will be generated:

- RO_000052326 (FCGR2A)
- RO_000059393 (FCGR3A)

The frequency of HLA-DRB1 positive assessments at baseline will be summarized.

Additionally, a summary of occurrence of AEs for <GENE> by genotype frequency for MS relapses, infections and IRRs will be generated for the same genetic markers.

Extended Periods

Above summaries and listings of safety assessments will be repeated for the ECP and OLE periods (including the SFU after withdrawal from ECP or OLE treatment).

These include but are not limited to;

- AEs leading to withdrawal from treatment and AEs leading to dose modifications
- Malignancies
- Serious infections
- Deaths
- IRRs by visit
- Summary of safety laboratory assessments, vital signs, antibody titers, immunoglobulins by visit
- Summary of treatment emergent ADAs
- Listing of pregnancy information

For OLE summaries, data will be presented as defined in OLE Safety Population.

Analyses pooling all safety data from the first dose of ocrelizumab (ECP and OLE periods) across multiple MS trials are described separately in a Global Safety Pooling SAP.

4.7.7 COVID-19 Analyses

In line with the guidance from the Sponsor and health authorities, several patient listings and summaries will be prepared in order to assess the impact of the Coronavirus Disease 2019 (COVID-19) pandemic on the study conduct and results. Summaries and listings will be prepared for major protocol deviations due to the pandemic, patients infected with COVID-19 and AEs & death associated with COVID-19. Further details will be provided in the COVID-19 CSR Appendix.

4.8 MISSING DATA

All methods for handling missing data and associated sensitivity analyses are described above, section by section, for each endpoint.

4.9 INTERIM ANALYSES

Updated efficacy and safety analysis of ECP and/or OLE are performed on a yearly basis after the CSR primary database lock. Regular safety updates are assessed in safety pooling analysis from multiple studies.

5. REFERENCES

Bohannon RW. Comfortable and maximum walking speed of adults aged 20−79 years: reference values and determinants. Age Ageing 1997;26:15−9.

Carpenter JR, Kenward MG. Missing data in clinical trials–a practical guide. NIHR Report. 2007. Available from: http://www.hta.nhs.uk/nihrmethodology/reports/1589.pdf.

Ebers GC. Natural history of primary progressive multiple sclerosis. Mult Scler 2004;10(Suppl 1):S8−13.

European Medicines Agency, Guideline on clinical investigation of medicinal products for the treatment of Multiple Sclerosis, 26 March 2015, EMA/CHMP/771815/2011, Rev. 2.

Fischer JS, Jak AJ, Kniker JE, Rudick RA,Cutter G: Multiple Sclerosis Functional Composite (MSFC) Administration and Scoring Manual 2002; National Multiple Sclerosis Society.

Hawker K, O'Connor P, Freedman MS, Calabresi PA, Antel J, Simon J, Hauser S, Waubant E, Vollmer T, Panitch h, Zhang J, Chin P, Smith CH. Rituximab in patients with primary progressive multiple sclerosis: results of a randomized double-blind placebo-controlled multicenter trial. 2009. Ann Neurol. 66:460-71.

Sahai H, Khurshid A. Confidence Intervals for the Mean of a Poisson Distribution: A Review. 1993. Biometrical J, 35: 857-67.

Weinshenker BG, Issa M, Baskerville J. Meta-analysis of the placebo-treated groups in clinical trials of progressive MS. Neurology 1996;46:1613−9.

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093

Universitätsspital Basel

Neurologische Klinik und Poliklinik

Standard Operating Procedure (SOP)

EDSS assessment check for the

Roche trials WA25046, WA21092 and WA21093

SD02

Attachments

Curriculum vitae of the USB Experts

Contract of the USB Experts in copy

Conflicts of interests of USB Experts

Version: 1.0

Target group: Roche, Neurostatus Systems GmbH, USB EDSS Expert Team

Function	Name, Company	Date	Signature
Author	[redacted] USB	[redacted]	
Approved	Prof. [redacted] GDTL, Roche		
Approved	Prof. [redacted] USB [redacted] of the Department of Neurology	[redacted]	
Approved	[redacted] Neurostatus Systems GmbH		
Approved	[redacted] Quality, USB		

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

Universitätsspital Basel

Neurologische Klinik und Poliklinik

Standard Operating Procedure (SOP)

EDSS assessment check for the

Roche trials WA25046, WA21092 and WA21093

SD02

Attachments

Curriculum vitae of the USB Experts

Contract of the USB Experts in copy

Conflicts of interests of USB Experts

Version: 1.0

Target group: Roche, Neurostatus Systems GmbH, USB EDSS Expert Team

Function	Name, Company	Date	Signature
Author	[redacted] USB		
Approved	Prof. [redacted] GDTL, Roche		
Approved	Prof. [redacted], USB [redacted] of the Department of Neurology		
Approved	[redacted] Neurostatus Systems GmbH		
Approved	[redacted] Quality, USB		

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

1. Aim

The aim of this SOP is to describe the process and the roles and responsibilities for the review of the Expanded Disability Status Scale (EDSS) assessments. In addition this SOP further outlines how to detect and handle inconsistent EDSS assessments and John Kurtzke's Functional Systems scores (FSS) in the Roche pivotal Multiple Sclerosis trials WA25046, WA21092 and WA21093 according to ICH guidelines.

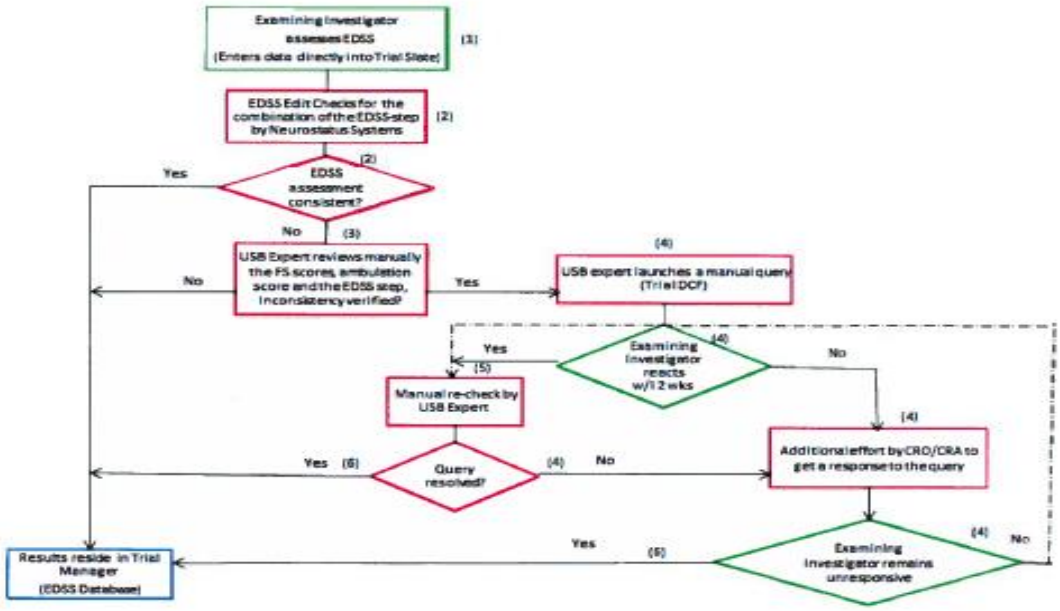
2. Background

Currently, John Kurtzke's Functional Systems and the EDSS are the most widely accepted clinical outcome measures for the evaluation of neurological impairment and disability in Multiple Sclerosis (MS) clinical trials. The determination of the EDSS step is primarily based on the individual scores of the 7 Functional Systems including visual, brainstem, pyramidal, cerebellar, sensory, bowel and bladder, and cerebral, as well as on the walking range and the assistance needed to ambulate (Kurtzke JF et al. 1955; Kurtzke JF 1983). During the past years standardized training programs and certification of EDSS evaluating physician (EDSS assessor) have been developed and introduced to improve inter-rater and intra-rater reliability. For the Roche trials WA25046, WA21092 and WA21093 the EDSS step is assessed by trained neurologists who refer to the Neurostatus definitions in the booklet version of 04/10.2. Between Jan 1, 2011 and Jan 31, 2012 University Hospital Basel (USB) Experts reviewed 1082 EDSS assessments rated by 267 examining investigators at 160 study sites participating in the Roche trials WA25046, WA21092 and WA21093. They found in 23% of the cases inconsistencies in the last step of the assessment, namely the combination of the Functional Systems and the ambulation scores to the final EDSS step.

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

3. Processes

3.1. Flow Chart on Query Resolution (Detail in 3.3)



3.2. Roles & Responsibilities

Responsible	Role
EDSS Evaluating Physician	Examining investigator / EDSS assessor who performs the neurological examination, documents the FS scores and assess EDSS steps. The examining Investigator will not be involved with any aspect of medical management of the patient and will not have access to patient data.
CRF Health (CRFHealth eCOA Solutions Company)	Provider of the TrialSlate and associated web portal (TrialManager). CRF Health is responsible for processing all data clarification forms (TrialDCFs) related to the EDSS assessment captured in the TrialManager.
Neurostatus Systems GmbH	Company responsible for the technical and administrative implementation of training and certification of physicians participating in projects using EDSS in multiple sclerosis. Neurostatus Systems is responsible to check for plausibility and inconsistencies of the EDSS assessments.
USB Expert	USB Experts are medical doctors working at the Department of Neurology of the USB specialized in the assessments of the EDSS. For contracts see attachment. Prof Kappos has the oversight of the Expert Review process. He will appoint a named person as one of the USB Experts who will act as a single point of contact for Roche.

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

	USB expert are reviewing EDSS assessment of the scoring sheet. The USB expert is responsible for the content of the TrialDCFs. The USB expert is responsible for being available via email / telephone in order to respond to questions by the EDSS evaluating physician.
Monitor/CRA	Clinical Research Associate employed by a Clinical Research Organisation (CRO). Monitor / CRA log into TrialManager to check for TrialDCFs and to notify sites that a response is required from the examining investigator.
Roche Study Management Team (SMT)	The Study Management Team (SMT) reviews unresolved TrialDCFs and decides to close or continue to follow up the outstanding TrialDCFs.

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

3.3. Reporting and Review Process in written form

Step 1: EDSS assessment by the examining investigator at the study sites

Functional system score, ambulation score and EDSS step (EDSS assessment) are assessed by examining investigators based on a standardized neurological examination using the Neurostatus Scoring booklet, version 04/10.2. These data are captured by using TrialSlate (an electronic data capturing device). Range checks are performed during data entry onto the device. The data are then transferred via LAN or Mobile network to TrialManager web portal (the database).

Step 2: Automated Consistency Check

Neurostatus Systems checks the data on TrialManager for plausibility and inconsistencies of the EDSS using automated consistency checks. The rules for these checks are given in the Neurostatus Scoring booklet, version 04/10.2. A scoring sheet consisting of the results of the EDSS assessments is generated by Neurostatus and uploaded to TrialManager. The scoring sheet will flag inconsistencies in the EDSS assessment.

If the scoring sheet doesn't identify inconsistencies, the assessment remains unchanged in TrialManager.

Step 3: Manual consistency check

The USB expert will review the scoring sheet with the EDSS assessments within 2 working days from upload to TrialManager.

The EDSS assessments with flagged inconsistencies in the scoring sheet are manually reviewed by the USB Expert.

- If after review by the USB expert, the flagged EDSS assessment is determined to be consistent then it will remain on TrialManager unchanged.
- If after review by the USB expert, the flagged EDSS assessments are confirmed to be inconsistent a manual query (TrialDCF) will be generated in TrialManager. The query in TrialManager will be reviewed and responded to by the examining investigator (See Step 4)

Step 4: TrialDCF

The TrialDCF are queries described in the study manual query process.

- The USB expert is responsible for the content of the TrialDCF

- CRF Health is responsible for processing all TrialDCFs related to the EDSS assessment captured in TrialManager.
- The Monitor / CRA will log into TrialManager at least once per week for any new TrialDCFs. If new TrialDCFs have been issued, sites will be notified by the monitor / CRA that a response to the TrialDCF is required by the examining investigator within two weeks via TrialManager.
- If the examining investigator does **not** respond to the initial TrialDCF within the two weeks as stated above, the monitor/CRA will contact the examining investigator to request to resolve the outstanding TrialDCF as per the agreed monitoring plan.
- The Examining investigator and/or monitor/CRA can ask for support by the USB expert via email or telephone to ask any question to enable resolution of the TrialDCF.
- Call service times are weekly for European countries on every Tuesday 3-4 p.m. CET, biweekly on Tuesdays 9-10 a.m. CET for Australia and Asia and 5-6 p.m. CET for the Americas. The content of these discussions during the calls will be documented and archived by the USB Expert.
- Detailed information on call in numbers and email information will be provided by the USB expert.
- All outstanding TrialDCFs will be followed up by the monitor/CRA as per the agreed monitoring plan. The Study Management Team (SMT) will review unresolved TrialDCFs (as defined in study Integrated Data Review Plan (iDRP)) and will decide to close or continue to follow up the outstanding TrialDCFs.
- Any TrialDCFs that are not answered/resolved will have their status changed as defined by CRFHealth (following authorisation from the Roche SMT during the course of the study (see Step 6).

Step 5: Manual re-check

All answered TrialDCFs will be re-checked by the same USB Expert who issued the initial TrialDCF. TrialManger will generate a report to notify the USB expert of the response from examining investigator at the site and he/she will review the response.

Step 6: Resolution of TrialDCF in TrialManager

If the TrialDCFs is resolved, the EDSS assessment will be considered final and CRFHealth will implement and verify the change in TrialManger

The final status of a TrialDCF will be either

1) "Resolved with change and closed" – TrialDCF is answered and revised data are entered by the examining investigator into TrialManager

2) "Resolved without change and closed" – TrialDCF is answered by the Examining Investigator who confirms his original assessment in the TrialDCF.

3) "Unresolved and closed" – TrialDCF is not answered and the SMT authorises to close due to unresponsiveness

In all cases the examining investigator remains the final decision maker on the EDSS assessment.

3.4 Handling of EDSS assessments checked prior to the implementation of the process outlined in this SOP

All previous EDSS assessments, that had gone through the previous automated TrialDCF process, whether they had been changed or not as a result of this process will go through this new revised Neurostatus Systems edit check process again and the process steps outlined in this current SOP 2-6 will be applied. Automated and new TrialDCF data will be stored in TrialManager.

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

4. Glossary & Definitions

ICH GCP guideline	International Conference on Harmonisation good clinical practice guidelines
CRF Health	CRFHealth eCOA Solutions Company
CRA	Clinical research associate
CRO	Contract research organisation
EDSS	Expanded Disability Status Scale
FSS	Functional system score
iDRP	Integrated Data Review Plan
MS	Multiple sclerosis
Neurostatus Systems GmbH	Company responsible for the technical and administrative implementation of training and certification of physicians participating in projects using EDSS in MS.
Principal Investigator	Lead site investigator
SMT	Study Management Team
TrialSlate	An electronic data capture device provided by CRFHealth
Trial DCF	Trial data clarification form (this resides in TrialManager)
TrialManager	Webportal Database which collects all questionnaires completed on the trialslate including EDSS assessments of the study patients
USB	University Hospital Basel
USB expert	Medical doctor who works at the University Hospital Basel (USB), specialized in the assessment of the EDSS and responsible for the quality control of inconsistencies, queries, teaching of the EDSS and certification process.

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

5. References

Kurtzke JF. A new scale for evaluating disability in multiple sclerosis. Neurology 1955;**5**:580-583.

Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;**33**(11):1444-1452.

Neurostatus, Version 04/10.2. www.Neurostatus.net

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

6. Amendments

Version	Section	Description of the amendments	Reason for the amendment
01		Initial version	

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

Supplements:

Curriculum vitae of the USB Experts

Contract of the USB Experts in copy

Conflicts of interests of USB Experts

Appendix 1
EDSS Assessment Check for the Roche Trials WA25046, WA21092, and WA21093 (cont.)

EDSS Assessment Check SD02

Universitätsspital Basel

Neurologische Klinik und Poliklinik

Example of an inconsistency: There is an inconsistency between the Ambulation Score and the EDSS Step based on the EDSS scoring rules booklet version 04/10.2.

STUDY NAME	WA25046
ASSESSMENT INFORMATION	
Assessment ID	320241500000000001494
Site ID	[redacted]
EDSS rater ID	
Subject ID	
Date of Examination	

SYNOPSIS	
1. Visual	1
2. Brainstem	0
3. Pyramidal	3
4. Cerebellar	3
5. Sensory	3
6. Bowel/Bladder	2
7. Cerebral	0

Ambulation Score	5
EDSS Step	5.5
Signature	

Signature Page for Study WA25046 (ORATORIO) SAP v5 - Published
System identifier: RIM-CLIN-472959

Approval Task	[redacted] Company Signatory 20-Mar-2023 15:24:56 GMT+0000