Official Title: A Phase IIIB, Multicenter, Randomized, Double-Blind, Controlled Study to Evaluate the Efficacy, Safety and Pharmacokinetics of a Higher Dose of Ocrelizumab in Adults with Relapsing Multiple Sclerosis

NCT Number: NCT04544436

Document Date: Statistical Analysis Plan Version 4: 18-Feb-2025

STATISTICAL ANALYSIS PLAN

STUDY TITLE: A PHASE IIIB MULTICENTER, RANDOMIZED, DOUBLE-BLIND, CONTROLLED STUDY TO EVALUATE THE EFFICACY, SAFETY AND PHARMACOKINETICS OF A HIGHER DOSE OF OCRELIZUMAB IN ADULTS WITH RELAPSING MULTIPLE SCLEROSIS

STUDY NUMBER: BN42082

STUDY NAME: MUSETTE

VERSION NUMBER: 4

ROCHE COMPOUND(S): Ocrelizumab (RO4964913)

EUDRACT NUMBER: 2020-000893-69

IND NUMBER: 100,593

NCT NUMBER: NCT04544436

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

STATISTICAL ANALYSIS PLAN APPROVAL

SPONSOR: F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS: Grenzacherstrasse 124
4070 Basel, Switzerland

DATE FINAL: See electronic date stamp on the last page of this document

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This Statistical Analysis Plan (SAP) was developed based on Roche SAP model document updated on 28 February 2022.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
4	See electronic date stamp on the last page of this document	Version 4.0, 21 December 2023
3	14 Oct 2024	Version 4.0, 21 December 2023
2	11 June 2024	Version 4.0, 21 December 2023
1	11 January 2023	Version 3.0, 27 October 2021

STASTISTICAL ANALYSIS PLAN AMENDMENT, VERSION 4: RATIONALE

Statistical Analysis Plan BN42082 has been updated primarily in the response to Food and Drug Administration (FDA) information request, ensuring alignment with the content and clarifications provided in the FDA response:

- **Clarifications and changes to the Sensitivity Analysis 1:** The condition that tied the conduct of the analysis to the number of Initial Disease Progression events for patients who withdrew from the study has been removed.

Additional minor changes have been made to improve clarity and consistency, including, but not limited to the following:

- Correction of the typo in the confirmation period (changed from 12 to 24) for the endpoint measuring the time from randomization to the first occurrence of 24-week confirmed $\geq$ 8-point increase in 12-Item Multiple Sclerosis Walking Scale (MSWS-12).
- One of the biomarker endpoints assessing blood B-cell levels by Fcγ Receptor 3A (FcγR3A) genotype affinity was duplicated under the pharmacodynamic objectives and has therefore been removed.
- The format of the output example provided in Appendix 1, supporting the impact assessment of the War in Ukraine on data integrity, has been modified to better reflect the potential impact of missing assessments on the primary estimand. This revision accounts for the fact that the time of event or censoring for the primary estimand is defined based on actual time rather than the scheduled visit label.

This amendment represents cumulative changes to the original Statistical Analysis Plan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDIX

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
9-HPT	9-hole peg test
ADA	anti-drug antibody
AE	adverse event
BAP	Biomarker Analysis Plan
cCDP	composite confirmed disability progression
cCDP12	12-week composite confirmed disability progression
cCDP24	24-week composite confirmed disability progression
cCDP48	48-week composite confirmed disability progression
CCOD	clinical cutoff date
CDP12	12-week confirmed disability progression based on the EDSS
CDP24	24-week confirmed disability progression based on the EDSS
CI	confidence interval
CI9-HPT12	12-week confirmed increase of ≥20% from baseline in 9-HPT score
CI9-HPT24	24-week confirmed increase of ≥20% from baseline in 9-HPT score
CIMSWS24	24-week confirmed ≥8-point increase in MSWS-12
CIT25FWT12	12-week confirmed increase of ≥20% from baseline in T25FWT score
CIT25FWT24	24-week confirmed increase of ≥20% from baseline in T25FWT score
C_{mean}	average concentration over the treatment period
COVID-19	coronavirus disease-2019
cPIRA12	cCDP12 independent of protocol-defined relapses
CRF	case report form
CWSDMT12	12-week confirmed ≥4-point worsening in Symbol Digit Modalities Test
DBT	double-blind treatment
DMT	disease-modifying therapy
eCRF	electronic Case Report Form
EDSS	expanded disability status scale
FAS	full analysis set
FDA	Food and Drug Administration
FSS	functional system scores
GCP	good clinical practice
HR	hazard ratio
iDMC	independent Data Monitoring Committee

IFN	interferon
IL-6	Interleukin 6
IMP	investigational medicinal product
IRR	infusion-related reaction
IV	intravenous
IxRS	interactive voice or web-based response system
MMRM	mixed model for repeated measures
MRI	magnetic resonance imaging
MS	multiple sclerosis
MSIS-29	Multiple Sclerosis Impact Scale-29
MSWS-12	12-Item Multiple Sclerosis Walking Scale
NfL	neurofilament light chain
NPMLE	non-parametric maximum likelihood estimate
OI	opportunistic infection
OLE	open-label extension
PCR	polymerase chain reaction
PD	pharmacodynamic
PK	pharmacokinetic
PPMS	primary progressive multiple sclerosis
PT	Preferred Term
RMS	relapsing multiple sclerosis
ROW	rest of world
RRMS	relapsing-remitting MS
SAE	serious adverse event
SAP	Statistical Analysis Plan
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SAS	safety analysis set
SDMT	Symbol Digit Modalities Test
SFU	safety follow-up
SMQ	Standardised MedDRA Queries
SOC	System Organ Class
T25FWT	timed 25-foot walk test

1. INTRODUCTION

Study BN42082 is a Phase III, multicenter, randomized, parallel-group, double-blind, controlled study to evaluate the efficacy and safety of a higher dose of ocrelizumab (1200 mg if body weight $<$75 kg and 1800 mg if bodyweight $\geq$75 kg) compared with the approved dose of ocrelizumab (600 mg) in patients with relapsing multiple sclerosis (RMS).

Even though there are many drugs currently available that target the acute inflammatory mechanisms associated with relapses and relapse associated worsening, to date, only ocrelizumab is indicated for primary progressive multiple sclerosis (PPMS) (note: ocrelizumab is only approved for active PPMS [aPPMS] in some countries). As a result, the salient feature of disability progression in all forms of multiple sclerosis (MS) remains to be further addressed, and treatments that can stop or delay MS disease progression represent a serious unmet medical need.

Ocrelizumab is a recombinant humanized, glycosylated, monoclonal IgG1 antibody that selectively targets and depletes CD20-expressing B cells, while preserving the capacity of B-cell reconstitution and preexisting humoral immunity. Based on the results of ocrelizumab Phase III studies in patient populations with RMS and PPMS, ocrelizumab was approved by the U.S. Food and Drug Administration (FDA) on 28 March 2017 for the treatment of adult patients with RMS and PPMS and by the European Medicines Agency on 12 January 2018 for patients with active RMS defined by clinical or imaging features and for patients with early PPMS in terms of disease duration and level of disability, and with imaging features characteristic of inflammatory activity.

The analyses described in the SAP will supersede those specified in the Protocol BN42082 for the purposes of regulatory filing.

Changes to the protocol-defined analyses are described in Section 4.7.

1.1 STUDY RATIONALE

In the Phase III pivotal studies with ocrelizumab 600 mg in patients with RMS (WA21092/WA21093) and PPMS (WA25046), a positive correlation between ocrelizumab exposure quartiles and the blood B-cell depletion level was observed. Moreover, higher exposure quartiles were associated with lower rates of confirmed disability progression. Namely, for the 12-week composite confirmed disability progression (cCDP12) endpoint, which includes time to progression measured by the Expanded Disability Status Scale (EDSS), Timed 25-Foot Walk Test (T25FWT) or 9-Hole Peg Test (9-HPT), those in the highest exposure quartile were estimated to have a hazard ratio (HR) of 0.50 (95% confidence interval (CI): [0.35, 0.72]) and 0.69 (95% CI: [0.52, 0.90]) versus interferon (IFN) and placebo for patients with RMS and PPMS, respectively. In comparison, those in the lowest exposure quartile were estimated to have a HR of 0.80 (95% CI: [0.58, 1.10]) and 0.84 (95% CI: [0.64, 1.11]) versus IFN and

placebo for patients with RMS and PPMS, respectively. In Study BN42082, the higher dose of ocrelizumab is predicted by the population pharmacokinetic (PK) model to bring 99% of patients to the highest exposure quartile where, by far, the best disability progression outcomes were observed in the Phase III pivotal studies.

At the same time, exposure-response relationships for safety parameters (serious adverse events [SAEs], serious infections, and infusion-related reactions [IRRs]) were also assessed, and no correlation between exposure quartiles and occurrence and grade of SAEs, serious infections, and IRRs was observed in patients receiving the ocrelizumab 600 mg intravenous (IV) regimen across both the RMS and PPMS populations. Furthermore, the maximum concentrations observed for ocrelizumab were not higher in patients that experienced at least one SAE compared with patients without SAEs, at the 600 mg dose. For both RMS and PPMS, the graphical analysis of the occurrence of SAE by System Organ Class (SOC) in patients treated with ocrelizumab showed no dependencies on exposure.

In the Phase II relapsing-remitting MS (RRMS) Study WA21493, patients were randomly assigned to either placebo, 600 mg, or 2000 mg of ocrelizumab, administered as two infusions of 300 mg or 1000 mg given 15 days apart. The 2000 mg dose of ocrelizumab was only given for Cycle 1, followed by single 1000 mg infusions for the subsequent dosing occasions. Overall, the study showed a favorable safety profile for both the 600 mg and 2000 mg doses of ocrelizumab. The PK analyses of the Phase II study data were also used to set the criterion for the highest ocrelizumab exposure to be tested in this study; an observed average concentration over the treatment period (C_{mean}) of 83 μg/mL was well tolerated in patients who received 2000 mg ocrelizumab. In Study BN42082, the higher dose of ocrelizumab is predicted by the population PK model to keep over 99% patients below the C_{mean} of 83 μg/mL.

In Study WA21493, pre-infusion treatment with an oral analgesic/antipyretic (e.g., acetaminophen), and an oral antihistamine (e.g., diphenhydramine) were proposed as recommendations only. As proven in the subsequent studies, the addition of oral antihistamine to methylprednisolone pretreatment for each infusion was associated with a lower incidence overall in IRRs compared with pretreatment with methylprednisolone alone. Therefore, mandatory antihistamine is also included in Study BN42082.

In summary, it is expected that higher doses of ocrelizumab will provide greater benefit to patients in further reducing disability progression, without compromising the safety profile of the drug.

1.2 OBJECTIVES, ENDPOINTS AND ESTIMANDS

Table 1 Primary and Secondary Objectives and Corresponding Estimands

Primary Objective	Estimand Definition
• To demonstrate the superiority of a higher dose of ocrelizumab over the approved dose of ocrelizumab	• **Population**: Participants with RMS as defined by the study inclusion and exclusion criteria (See Sections 4.1.1 and 4.1.2 of the protocol) • **Endpoint**: time from randomization to the first occurrence of cCDP12 according to at least one of the following three criteria: o 12-week confirmed disability progression (CDP12), defined as a 12-week confirmed increase from baseline in EDSS score of ≥ 1.0 point in patients with a baseline EDSS score of ≤ 5.5 or a 12-week confirmed increase ≥ 0.5 points in patients with a baseline EDSS score of > 5.5, or o 12-week confirmed increase in T25FWT (CIT25FWT12), defined as a 12-week confirmed increase of ≥ 20% from baseline in T25FWT score, or o 12-week confirmed increase in 9-HPT (CI9-HPT12), defined as a 12-week confirmed increase of ≥ 20% from baseline in time to complete the 9-HPT score. • **Treatments**: o Experimental arm: 1200 mg if body weight at baseline < 75 kg and 1800 mg if bodyweight at baseline ≥ 75 kg every 24 weeks o Control arm: 600 mg every 24 weeks • **Intercurrent events and handling strategies**: –**Early discontinuation from study treatment:** Treatment-policy strategy –**Initiation of an MS disease-modifying therapy (DMT) medication**: Treatment-policy strategy –**Pregnancy**: Treatment-policy strategy. –**Death**: Composite strategy For more details, see Section 4.2. • **Population-level summary**: HR for time to onset of cCDP12

Secondary Objectives	Estimand/Endpoint Definition
• To demonstrate superiority of a higher dose of ocrelizumab over the approved dose of ocrelizumab	• **Population**: as defined above • **Endpoints:** ○ Time from randomization to the first occurrence of 24-week composite confirmed disability progression (cCDP24) ○ Time from randomization to the first occurrence of 48-week composite confirmed disability progression (cCDP48) ○ Time from randomization to the first occurrence of 12-week confirmed disability progression (CDP12) ○ Time from randomization to the first occurrence of 12-week confirmed increase in T25FWT (CIT25FWT12) ○ Time from randomization to the first occurrence of cCDP12 independent of protocol-defined relapses (cPIRA12) ○ Time from randomization to the first occurrence of 12-week confirmed ≥ 4-point worsening in SDMT (CWSDMT12) – Time from randomization to the first occurrence of 24-week confirmed ≥ 8-point increase in MSWS-12 (CIMSWS24) • **Treatments:** as defined above • **Intercurrent events and handling strategies:** as defined above • **Population-level summary:** hazard ratio for time to onset of endpoints specified above See Section 4.2.2 for more details.

• To demonstrate superiority of a higher dose of ocrelizumab over the approved dose of ocrelizumab	• **Population**: as defined above • **Endpoint:** Annual rate of percent change from baseline in total brain volume • **Treatments:** as defined above • **Intercurrent events and handling strategies**: –**Early discontinuation from study treatment:** Hypothetical strategy –**Initiation of an MS disease-modifying therapy (DMT) medication**: Hypothetical strategy –**Pregnancy**: Treatment-policy strategy. –**Death**: Hypothetical strategy • **Population-level summary measure**: mean difference between treatment arms in the annual rates of percent change from baseline in total brain volume
• *To demonstrate that both the higher dose and standard dose of ocrelizumab can significantly reduce neurofilament light chain (NfL) from baseline*	• **Population**: as defined above. • **Endpoint:** Change from baseline in neurofilament light chain (NfL), i.e. ratio to baseline, at Week 48 • **Treatments:** as defined above • **Intercurrent events and handling strategies**: –**Early discontinuation from study treatment:** Hypothetical strategy –**Initiation of an MS disease-modifying therapy (DMT) medication**: Hypothetical strategy –**Pregnancy**: Treatment-policy strategy. • –**Death**: Hypothetical strategy • **Population-level summary measure**: mean change in NfL from baseline, i.e. ratio to baseline, at Week 48 within each treatment arm

9–HPT = 9–hole peg test; cCDP12 = 12-week composite confirmed disability progression; cCDP24 = 24-week composite confirmed disability progression; CDP12 = 12-week confirmed disability progression; CDP24 = 24-week confirmed disability progression; CI9-HPT12 = 12-week confirmed increase of ≥ 20% from baseline in 9-HPT score; CIT25FWT12 = 12-week confirmed increase of ≥ 20% from baseline in T25FWT score; CIMSWS24 = 24-week confirmed ≥ 8 point increase in MSWS-12; cPIRA12 = cCDP12 independent of protocol-defined relapses; CWSDMT12 = Time to 12-week confirmed 4-point worsening in SDMT; DMT = disease-modifying therapy; EDSS = expanded disability status scale; HR = hazard ratio; MS = multiple sclerosis; MSWS-12 = 12-item Multiple Sclerosis Walking Scale; NfL= neurofilament light chain; RMS = relapsing multiple sclerosis; SDMT = Symbol Digit Modalities Test; T25FWT = timed 25-foot walk test.

Table 2 Exploratory Efficacy Objectives and Endpoints

Exploratory Objectives	Endpoints
The exploratory efficacy objective is to evaluate the efficacy of a higher dose of ocrelizumab compared with the approved dose of ocrelizumab	• Time to onset of cPIRA24 and its individual components • Time to onset of CIT25FWT24 • Time to onset of CDP24 • Time to onset of CI9-HPT12 • Time to onset of CI9-HPT24 • Time from randomization to the first occurrence of 12-week confirmed ≥ 8-point worsening in SDMT • Time to onset of cCDP12 or 12-week confirmed ≥ 4-point worsening in SDMT • Time to onset of cCDP12 or 12-week confirmed ≥ 8-point worsening in SDMT • Change from baseline in EDSS score at Week 48 and 120 • The following patient-reported outcomes: o Change from baseline in the Multiple Sclerosis Impact Scale-29 (MSIS-29) physical scale at Week 48 and 120 o Change from baseline in MSIS-29 psychological scale at Week 48 and 120 o Change from baseline in Quality of Life in Neurological Disorders (Neuro-QOL) Upper Extremity Function Form at Week 48 and 120 o Change from baseline in Modified Fatigue Impact Scale at Week 48 and 120

Exploratory Objectives	Endpoints
	○ Proportion of patients with no change, improvement or worsening in overall severity in patient global impression of severity at Week 48 and 120 ○ Proportion of patients with no change, improvement or worsening in overall MS symptoms measured with patient global impression of change at Week 48 and 120 ○ Proportion of patients with no change, improvement or worsening in tasks involving arms/hands measured with patient global impression of change of the upper limb function at Week 48 and 120 • Annualized protocol-defined relapse rate • Time to onset of 12-week confirmed protocol-defined relapse associated worsening and individual components • Time to onset of progression in cCDP12 individual components independent of protocol-defined relapses • Total number of new T1-hypo-intense lesions (black holes) • Annual rate of change from baseline in total radius of T1-hypo-intense lesions (black holes) • Annual rate of percentage change from baseline in thalamic volume • Annual rate of percentage change in volume of spinal cord (upper part of the spine) • Total number of new or enlarging T2 lesions per MRI scan over the 120-week treatment period and at each scheduled visit • Total number of T1Gd+ lesions over the 120-week treatment period and at each scheduled visit

cCDP12 = 12-week composite confirmed disability progression; CDP24 = 24-week confirmed disability progression; CI9-HPT12 = 12-week confirmed increase of ≥20% from baseline in 9-HPT score; CI9-HPT24 = 24-week confirmed increase of ≥20% from baseline in 9-HPT score;CIT25FWT24 = 24-week confirmed increase of ≥20% from baseline in T25FWT score; cPIRA24 = cCDP24 independent of protocol-defined relapses; EDSS = expanded disability status scale; MRI = magnetic resonance imaging; MS = multiple sclerosis; MSIS-29 = Multiple Sclerosis Impact Scale-29; SDMT = Symbol Digit Modalities Test; T25FWT = timed 25 foot walk test.

Table 3 Other Objectives and Endpoints

Objectives	Endpoints
The safety objective is to evaluate the safety profile of a higher dose of ocrelizumab compared with the approved dose of ocrelizumab as well as the overall safety profile and safety profile by treatment arm over time	• Incidence and severity of adverse events, with severity determined according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) Version 5.0 • Change from baseline in clinical laboratory test results (including hematology, chemistry, and Ig levels) • Change from baseline in vital signs (including systolic and diastolic blood pressure, and pulse rate) following study treatment administration
The PK objective is to assess the exposure to ocrelizumab in serum in all patients in both study arms	• Serum concentration of ocrelizumab at specified time points, and derived PK parameters via the population PK approach
The pharmacodynamic (PD) objective is to characterize the ocrelizumab PD profile	• B-cell levels in blood (including comparing the degree of B-cell depletion between the doses) • Proportion of patients achieving 5 or less B-cells per microliter of blood • Proportion of patients achieving 0.4 or less B-cells per microliter of blood
The immunogenicity objective is to evaluate the immune response to ocrelizumab	• Prevalence of anti-drug antibodies (ADAs) at baseline and incidence of ADAs during the study
The exploratory PK objectives are to evaluate a potential relationship between drug exposure and the efficacy and safety of ocrelizumab	• Correlation of ocrelizumab serum concentration with efficacy endpoints • Correlation of ocrelizumab serum concentration with safety endpoints
The exploratory immunogenicity objective is to characterize the relationship between ADA status and the following endpoints	• Efficacy, safety, PD or PK endpoints

Objectives	Endpoints
The exploratory biomarker objectives are to identify biomarkers that are predictive of response to a higher dose of ocrelizumab (i.e., predictive biomarkers), are early surrogates of efficacy, are associated with progression to a more severe disease state (i.e., prognostic biomarkers), are associated with acquired resistance to ocrelizumab, are associated with susceptibility to developing adverse events or can lead to improved adverse event monitoring or investigation (i.e., safety biomarkers), can provide evidence of ocrelizumab activity (i.e., PD biomarkers), or can increase the knowledge and understanding of disease biology and drug safety. The following biomarker analyses may be implemented	• Levels of soluble biomarkers including but not limited to NfL and/or IL-6 in blood (plasma and/or serum) • Levels of B or T cell subsets in blood, including but not limited to CD19 + IgD, CD27, CD38, CD4, CD8, CD3, parameters to identify B or T naive, memory and/or B plasma blast/plasma cell subsets, DNA genotype of patients to include but not be limited to Fcγ R3A and human leukocyte antigen genotype
The exploratory health status objective is to evaluate health status of patients treated with a higher dose of ocrelizumab and with the approved dose of ocrelizumab	• Change from baseline at each scheduled visit up to and including Week 120 in EuroQol 5-Dimension instrument (5-level version; EQ-5D-5L)
The exploratory vaccine immune response objective is to investigate the effect of the higher and approved dose of ocrelizumab on antibody and T-cell responses in patients administered with an approved or authorized COVID-19 vaccine	• SARS-CoV-2 antibody titers • SARS-CoV-2 T-cell responses

ADA = anti-drug antibody; COVID-19 = coronavirus disease-2019; DNA = deoxyribonucleic acid; IL-6 = Interleukin 6; NCI CTCAE = National Cancer Institute Common Terminology Criteria for Adverse Events; NfL = neurofilament light chain; PD = Pharmacodynamic; PK = pharmacokinetic; SARS-CoV-2 = severe acute respiratory syndrome coronavirus 2.

1.3 STUDY DESIGN

Study BN42082 is a Phase IIIb, randomized, double-blind, controlled, parallel group, multicenter study to evaluate efficacy, safety and pharmacokinetics of a higher dose of ocrelizumab (1200 mg [patient's body weight $<$ 75 kg at baseline] or 1800 mg [patient's body weight $\geq$ 75 kg at baseline]) per IV infusion every 24 weeks in patients with RMS, in comparison to the approved 600 mg dose of ocrelizumab. Patients will be treated for a minimum of 120 weeks, representing at least five 24-week treatment doses, and until approximately 205 confirmed composite confirmed disability progression (cCDP) events have occurred.

This study will consist of the following phases: screening, double-blind treatment (DBT) phase, open-label extension (OLE) phase, safety follow-up (SFU), and B cell monitoring.

The primary analysis of the study will be performed when 205 cCDP events have occurred and the last enrolled patient completes at least 120 weeks of study treatment in DBT. All patients will stay on their randomized treatment until the primary efficacy analyses.

This document describes the primary analysis. At the time of the primary analysis, the open-label treatment period will not have started. The analysis of data from the open-label treatment phase is not described in this document.

1.3.1 Treatment Assignment

This is a randomized, double-blind study. After initial written informed consent has been obtained, all screening and baseline procedures and assessments have been completed, and eligibility has been established for a patient, the patient will be assigned an identification number and treatment assignment by an interactive voice or web-based response system (IxRS). Patients will be randomly assigned to one of two treatment arms: higher dose or approved dose of ocrelizumab. Randomization will occur in a 2:1 ratio (higher dose to approved dose, respectively) through use of a permuted-block randomization method to ensure a balanced assignment to each treatment arm.

Randomization will be stratified according to the following criteria:

- Weight at baseline ($<$ 75 kg vs. $\geq$75 kg)
- Region (United States vs. Rest of World [ROW])
- EDSS ($<$4 vs. $\geq$4)
- Age at screening ($\leq$45 years vs. $>$45 years)

1.3.2 Blinding

Study site personnel and patients will be blinded to treatment assignment during the study. The Sponsor and its agents will also be blinded to treatment assignment, with the exception of individuals who require access to patient treatment assignments to fulfill their job roles during a clinical trial. These roles will be documented in a separate study document and may include, but may not be limited to, the unblinding group responsible, clinical supply chain managers, sample handling staff, operational assay group personnel, IxRS service provider, and independent Data Monitoring Committee (iDMC) members. The handling of the investigational medicinal product (IMP) by the pharmacist as well as the process to maintain the blind is described in the pharmacy manual. If unblinding is necessary for a medical emergency (e.g., in the case of an SAE for which patient management might be affected by knowledge of treatment assignment), the investigator will be able to break the treatment code by contacting the IxRS. The investigator is not required to contact the Medical Monitor prior to breaking the treatment

code; however, the treatment code should not be broken except in emergency situations.

If the investigator wishes to know the treatment assignment for any reason other than a medical emergency, he or she should contact the Medical Monitor directly. The investigator should document and provide an explanation for any non-emergency unblinding. If the Medical Monitor agrees to unblinding of treatment assignment, the investigator will be able to break the treatment code by contacting the IxRS.

Any unblinding at the investigating site will be documented in the study report with the date, reason for identifying the assigned treatment/drug dose and name of the persons who requested the unblinding.

As per health authority reporting requirements, the Sponsor's Drug Safety representative will break the treatment code for all serious, unexpected suspected adverse reactions (see Section 5.7 of the protocol) that are considered by the investigator or Sponsor to be related to study drug. The patient may continue to receive treatment, and the investigator, patient, and Sponsor personnel, with the exception of the Drug Safety representative and personnel who must have access to patient treatment assignments to fulfill their roles (as defined above), will remain blinded to treatment assignment.

To maintain integrity of the trial results and to prevent potential unblinding of the assigned arm during the DBT phase as a result of AEs or changes to laboratory results, the following additional measures will be implemented until the time of the primary analysis:

- To prevent potential unblinding as a result of AEs or laboratory changes, a "dual assessor" approach will be used to evaluate efficacy and safety. Each site will have two blinded investigators: a Principal or Treating Investigator, and a rating or Examining Investigator.

  The Treating Investigator will be the safety assessor and should be a neurologist with experience in the care of patients with MS. The Treating Investigator will have access to safety data only and will make all treatment decisions based on the patient’s clinical response and laboratory findings.

  The Examining Investigator will be the efficacy assessor and should be a neurologist or other qualified health care practitioner trained and certified (where required) in administering and scoring the EDSS, T25FWT, 9-HPT, Functional System Scores (FSS), and Symbol Digit Modalities Test (SDMT). The Examining Investigator (or their designee) will assess the EDSS, T25FWT, 9-HPT, and SDMT. During the DBT phase, the Examining Investigator and their qualified designees (if applicable) will not be involved with any aspect of the medical management of the patient and will not be allowed access to patient data. They will also not discuss any treatment aspects (i.e., AEs) with the patient.

The Treating Investigator and the Examining Investigator will not be allowed to switch roles. During the DBT, an investigator/site staff at a single site must not be a Treating Investigator for some patients and an Examining Investigator for others.

The Treating Investigator will confirm the screening and baseline eligibility criteria for EDSS, 9-HPT, and T25FWT by accessing the relevant reports.

- Patient education: During the DBT phase, prior to being examined by the Examining Investigator, patients should be instructed not to discuss with the Examining Investigator what (if any) AEs they may be experiencing.
- Blinded, central (Magnetic resonance imaging) MRI assessments: A blinded, central MRI reader will assess all MRI scans performed during the study. During the DBT phase, only the local radiologist/technician at the investigational site who is assigned to this study may have access to the MRI scans. Of note, baseline MRI scan will be used for the assessment of patient eligibility, and therefore it will not be blinded. All MRI scans will also be reviewed locally by a radiologist for safety, and the MRI scan report containing only non-MS pathology will be provided to the Treating Investigator (refer also to Section 4.5.7 of the protocol).
- Blinding of laboratory parameters: Selected laboratory parameters that may lead to unblinding of the treatment assignment, such as flow cytometry assessment of cell counts (including CD19+ cells, lymphocyte count, absolute neutrophil counts [ANC], CD4) and Ig levels will be blinded in all patients until the start of OLE, which will occur after the primary analysis is complete. To ensure patient safety during the study and to allow for assessments of the re-treatment criteria, a central laboratory will provide study investigators and the Medical Monitor(s) with reflex messages triggered by abnormal blinded laboratory results and will be instructed to suspend further treatment with study drug until the patient becomes eligible for ocrelizumab re-treatment. Investigators will be notified of their patient's abnormal laboratory test results (via reflex messages). Refer to the laboratory manual for additional information.
- Treatment allocation will remain blinded at least until the database lock for the primary analysis.

Of note, the blinding procedures are not necessary during the OLE phase. It is, however, recommended that the same Examining Investigator continues to perform the assessments throughout the OLE phase as in the DBT phase.

1.3.3 Independent Review Facility

Some efficacy endpoints (e.g., total brain volume, thalamic volume, number and volume of new T1-hypo-intense lesions, volume of spinal cord, total number of new and enlarging T2 lesions, and T1Gd+ lesions over the 120-week treatment period and at each scheduled visit) are determined using MRI scans. The 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. Of note, baseline MRI scan will be used for the assessment of patient eligibility, and therefore it will not be blinded. 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.

1.3.4 Cleaning Process for EDSS, T25FWT, 9-HPT and SDMT

The primary efficacy endpoint will be derived from the EDSS, T25FWT and 9-HPT values recorded at any visits. The key secondary efficacy endpoint for SDMT is derived from the SDMT values recorded at any visit. These assessments are administered by an Examining Investigator (not the Treating Investigator), entered into an electronic device, and transferred to a central database for data-cleaning activities.

1.3.5 Data Monitoring

An independent Data Monitoring Committee (iDMC) will be employed to monitor and evaluate patient safety throughout the study, until the primary analysis is performed. Monitoring details are described in the iDMC Charter.

1.3.6 Good Clinical Practice Breach

All serious good clinical practice (GCP) breaches will be reported in the clinical study report (CSR).

If a GCP breach prevents the sponsor from verifying the validity of the data from a particular site, then the resulting data from that site will not be used in the data analyses described in Section 4.

2. STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATION

2.1 STATISTICAL HYPOTHESES

The hypotheses to be tested for the primary estimand are:

- H0 (null hypothesis): There is no difference in the time to onset of cCDP12 between the higher and approved dose of ocrelizumab arms.
- H1 (alternative hypothesis): There is a difference in the time to onset of cCDP12 between the higher and approved dose of ocrelizumab arms.

This will be tested at the 5% significance level ($\alpha = 0.05$) against two-sided alternatives.

Null and alternative hypothesis of similar form will be tested for all secondary efficacy endpoints, i.e., there is no difference (null hypothesis) vs there is a difference (alternative hypothesis) in endpoint between the higher and approved dose of ocrelizumab arms, except for NfL endpoints within each study arm where the hypotheses to be tested are:

- H0 (null hypothesis): There is no change in NfL from baseline at Week 48 within the higher or approved dose of ocrelizumab arms.

- H1 (alternative hypothesis): There is a change in NfL from baseline at Week 48 within the higher or approved dose of ocrelizumab arms.

The secondary efficacy estimands will be tested in the hierarchical order listed below, if the primary estimand and each preceding estimand have reached the significance level of 0.05:

- Time to onset of 24-week composite confirmed disability progression (cCDP24)
- Time to onset of 12-week confirmed increase in T25FWT (CIT25FWT12)
- Time to onset of cCDP12 independent of protocol-defined relapses (cPIRA12)
- Change from baseline in NfL at Week 48 for patients assigned to OCR HD arm
- Change in NfL from baseline at Week 48 for patients assigned to OCR approved dose arm
- Time to onset of 12-week confirmed disability progression (CDP12)
- Time to onset of 48-week composite confirmed disability progression (cCDP48)
- Time to onset of 12-week confirmed ≥ 4-point worsening in Symbol Digit Modalities Test (CWSDMT12)
- Time to onset of 24-week confirmed ≥ 8-point increase in 12-Item Multiple Sclerosis Walking Scale (CIMSWS24)
- Annual rate of percent change from baseline in total brain volume

The full description of the estimand attributes for the secondary estimands is provided in Table 1.

2.2 SAMPLE SIZE DETERMINATION

The sample size was determined based on data from Studies WA21092 and WA21093 (Opera I and Opera II, respectively) and Kappos et al. (2018). A two-group test of equal exponential survival curves was used to determine the sample size for the time to onset of cCDP12. With a 2:1 randomization ratio between the higher and approved dose treatment arm and assuming recruitment of 390 patients per year, 5% annual dropout rate, a control group progression rate of 21.5% and 5% two-sided Type I error, it is predicted that a sample size of approximately 786 (524:262) patients will be required to achieve 80% power to detect a HR of 0.66 based on 205 cCDP12 events. Therefore, the DBT phase will run until at least 205 cCDP12 events are observed.

Between November 2020 and December 2021, a much faster recruitment rate was observed than anticipated and 798 patients were randomized. Of the 798 randomized patients, 116 and 135 patients were participating at sites in Ukraine and Russia, respectively. To mitigate a potential withdrawal of patients or loss of data from Ukraine and Russia, due to the current conflict in the Ukraine, starting in February 2022, 67 additional patients were randomized from April 2022 to August 2022 resulting in a total sample size of 865 patients (10% more than the planned sample size in the absence of the war) for this study.

Supplementary analyses have been proposed in Section 4.2.4 to investigate the impact of the Russia-Ukraine war on the primary analysis. Furthermore, analyses to assess the impact of the war on study conduct have been proposed in Section 4.6.3.

2.3 ANALYSIS TIMING

The clinical cutoff date (CCOD) will be approximately when the last enrolled patient completes at least 120 weeks of study treatment, provided that the total number of cCDP12 events for the primary efficacy analysis is approximately 205. As of the present time, 205 cCDP events have been observed. Therefore, the CCOD will take place when the last enrolled patient completes at least 120 weeks of study treatment, which is currently scheduled to be in December 2024.

Unblinding of the Sponsor will occur several weeks following the CCOD when all data up to the CCOD are cleaned and the database is locked for the primary analysis (called the primary database lock in the remainder of the document). The primary analysis will include all data up to the CCOD for both safety and efficacy endpoints, unless stated otherwise.

The primary and secondary efficacy endpoints that are not planned to be analyzed with respect to a specific time point will use all available data in the database at the time of primary database lock. The other endpoints that are planned to be analyzed with respect to a specific time point (e.g., change from baseline to Week 48 or 120) will use data collected for each patient only up to and including the pre-specified time point. Any data collected for patients beyond that time point will not be included in the endpoint derivation or analysis.

At the end of the study an additional analysis comprising all data collected including the OLE will be performed. These analyses will be documented in a separate SAP.

3. ANALYSIS SETS

All efficacy endpoints will be analyzed using the full analysis set (FAS), which is described in Table 4. Table 4 also covers analysis sets relevant to the PK, immunogenicity, biomarker and safety objectives of this study.

Table 4 Participant Analysis Sets

Participant Analysis Set	Description
Full analysis set	All randomized participants: participants will be included in the analyses according to the treatment they were assigned.
Immunogenicity analysis set	All participants with at least one ADA assessment; participants will be grouped according to treatment received.
Pharmacokinetic analysis set	All participants who have measurable concentrations of ocrelizumab unless major protocol deviations or unavailability of information (e.g., exact blood sampling time) occurred or if data are unavailable, not plausible, or incomplete which may interfere with PK evaluation. Excluded cases will be documented together with the reason for exclusion.
Safety analysis set	All participants who received at least one infusion (partial or complete) of study drug; participants will be included in the analyses according to the treatment they received.

ADA = anti-drug antibody; PK = pharmacokinetic.

4. STATISTICAL ANALYSES

4.1 GENERAL CONSIDERATIONS

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.

All primary and secondary efficacy endpoints will be stratified by region, age at screening, baseline EDSS and baseline body weight for analysis. The regional stratifications are United States and ROW. The age categories are ≤ 45 years and > 45 years. The EDSS categories are < 4.0 and ≥ 4.0. The body weight categories are < 75 kg and ≥ 75 kg.

The baseline for the primary endpoint is the date of randomization. The baseline value of the endpoint, e.g., EDSS, T25FWT and 9-HPT values, is the most recent value up to and including the date of randomization but before first infusion of study drug. A subject with delayed dosing may receive the first dose 2 weeks after the baseline visit; if an EDSS, T25FW or 9-HPT score is recorded between the randomization date and the date of the first dose, this value will be considered as a post-baseline assessment.

FAS will be used in the analysis of the primary and secondary estimands, and exploratory efficacy endpoints. Safety analysis set (SAS) will be used to summarize safety data. For analysis purposes, the time of treatment discontinuation will be defined as 24 weeks after the last dose prior to the date of treatment discontinuation recorded in the electronic Case Report Form (eCRF).

4.2 PRIMARY ESTIMAND ANALYSIS

Hypothesis test for the primary estimand is described in Section 2.1.

Five attributes of the primary estimand are provided in Table 1. Further details on the definition of the Endpoint attribute are provided in Section 4.2.1. Information on primary estimator, intercurrent events and their handling as well as the handling of missing data is described in Section 4.2.2, Section 4.2.3 and Section 4.2.4 describe sensitivity and supplementary analysis for the primary estimand.

4.2.1 Definition of Primary Endpoint

The primary efficacy endpoint is time to onset of cCDP12, defined as the first occurrence of a CDP event according to at least one of the following three criteria:

- **C1**: CDP12, defined as a 12-week confirmed increase from baseline in EDSS score of ≥ 1.0 point in patients with a baseline EDSS score of ≤ 5.5 or a 12-week confirmed increase ≥ 0.5 points in patients with a baseline EDSS score of > 5.5, or
- **C2**: CIT25FWT12, defined as a 12-week confirmed increase of $\geq 20\%$ from baseline in T25FWT score, or
- **C3**: CI9-HPT12, defined as a 12-week confirmed increase of $\geq 20\%$ from baseline in time to complete the 9-HPT score.

Four conditions must be fulfilled to satisfy the 12-weeks confirmation criteria for disability progression based on C1, C2 and C3:

1. The confirmation (of disability progression) visit must be at least 12 weeks (84 days) after the initial disability progression.
2. The confirmation visit must be either a scheduled or treatment discontinuation visit.
3. Assessments within 30 days after the onset of a protocol-defined relapse cannot be used for confirmation of the progression.
4. All EDSS, T25FWT or 9-HPT assessments between the initial disability progression and the confirmation visit should also fulfill the requirements of disability progression, i.e., ≥ 1.0 point increase in baseline EDSS if baseline EDSS ≤ 5.5 or ≥ 0.5 increase if baseline EDSS > 5.5, or $\geq 20\%$ increase in T25FWT or 9-HPT from baseline.

If any of the four criteria listed above are not satisfied, then the initial disability progression is not a CDP. In criteria 3, a **protocol-defined relapse** is defined as the occurrence of new or worsening neurological symptoms attributable to MS and immediately preceded by a relatively stable or improving neurological state of at least 30 days. Symptoms must persist for at least 24 hours and should not be attributable to confounding clinical factors (e.g., fever, infection, injury, adverse reactions to concomitant medications). The new or worsening neurological symptoms must be

accompanied by objective neurological worsening consistent with an increase of at least one of the following:

- Half a step (0.5 point) on the EDSS
- Two points on one of the selected FSS as listed below
- One point on two or more of the selected FSS as listed below

The change must affect the following selected FSS: pyramidal, ambulation, cerebellar, brainstem, sensory, or visual. Episodic spasms, sexual dysfunction, fatigue, mood change, or bladder or bowel urgency or incontinence will not suffice to establish a relapse.

An independent Examining Investigator at each study site will assess EDSS, T25FWT and 9-HPT 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, and at withdrawal-from-treatment and end-of-study visits. Additional assessments for individual patients may be requested between visits (i.e., during an MS relapse). Initial disability progression can occur at any visit, including unscheduled visits.

The Examining Investigator is not the physician responsible for the patient care (the Treating Investigator).

The EDSS (2011 Version 04/10.2) is based on a standard neurological examination, incorporating functional systems (visual, brainstem, pyramidal, cerebellar, sensory, bowel and bladder, and cerebral [or mental]) that are rated and then scored as a FSS, and ambulation, which is scored as ambulation score. Each FSS is an ordinal clinical rating scale ranging from 0 to 5 or 6 and an ambulation score that is rated from 0 to 16. These ratings are then used in conjunction with observations, as well as information, concerning ambulation and use of assistive devices to determine the total EDSS score. The total EDSS score ranges from 0 (normal) to 10.0 (death) in 0.5-point steps. All FSS, ambulation score, and total EDSS scores will be captured electronically and the Neurostatus-eEDSS definitions and calculating algorithms will be used (D'Souza et al. 2017). Under this approach, all subscores are required (meaning they are compulsory and data entry is required) and all subscores always contribute to the scoring algorithm with the following exceptions:

- Cerebral FS
  - Fatigue

The following subscores are always optional (meaning data entry is not required), but if a value for one side (right or left) is entered, the other is also required. In other words, the user must either enter both right and left values, or neither. When both right and left values are entered, they always contribute to the scoring algorithm:

- Visual FS
  - Disc Pallor
- Pyramidal FS
  - Palmomental Reflex

The following subscores are always optional, and they never contribute to the scoring algorithm:

- Pyramidal FS
  - Pronator Drift – pronation
  - Pronator Drift - downward drift
  - Position Test – sinking
  - Walking on Heels
  - Walking on Toes
  - Hopping on One Foot
  - If able to lift only one leg at a time: (grade in degrees)
- Sensory FS
  - Lhermitte's Sign
  - Paraesthesiae UE
  - Paraesthesiae trunk
  - Paraesthesiae LE
- Bowel & Bladder FS
  - Sexual Dysfunction
- Ambulation
  - Distance reported by the patient
  - Time reported by the patient

The T25FWT test is a performance measure used to assess walking speed based on a timed 25-foot walk. The patient is directed to start at one end of a clearly marked 25-foot course and is instructed to walk 25 feet as quickly and safely as possible. The Examining Investigator will time the patient from the start of the walk to the end of the 25 feet. The task is immediately administered again by having the patient walk back the same distance. The score for the T25FWT is the average of the two completed trials.

Patients may use assistive devices (e.g., cane, crutch, or rollator) when performing the task. The Examining Investigator should assess which assistive device is necessary and try to use the same assistive device at each study visit, whenever possible. Non-wheeled walkers should not be used. Circumstances that may affect the patient's performance, or if the patient cannot complete the T25FWT twice, should be reported.

If the test results for T25FWT 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. Very low values will be considered to be outliers. For outliers, the following rules will be applied. Values below the lower bound will be treated as missing, and the imputation rule will be applied as defined above. The lower bound is 2.2 seconds (Bohannon 1997).

The 9-HPT is a performance measure to assess upper limb function. It consists of a container containing nine pegs and a wood or plastic block containing nine empty holes. The patient is to pick up each of the nine pegs one at a time and as quickly as possible place them in the nine holes. Once all the pegs are in the holes, the patient is to remove them again one at a time as quickly as possible and replace them into the container. The total time to complete the task is recorded. Both the dominant and non-dominant hands are tested twice (two consecutive trials of the dominant hand, followed immediately by two consecutive trials of the non-dominant hand). The score for the 9-HPT is an average of the four trials. The two trials for each hand are averaged, converted to the reciprocal of the mean time for each hand, and then the two reciprocals are averaged and back-transformed to the original scale (i.e., by taking another reciprocal).

If the test results for 9-HPT are not available due to a "physical limitation," the maximum possible value for the scale (300 seconds) will be imputed. If one of the two trials for the dominant and non-dominant hand 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. Very low values will be considered to be outliers. For outliers, the following rules will be applied. Values below the lower bound will be treated as missing, and the imputation rule will be applied as defined above. For 9-HPT, the lower bound is 10 seconds (Oxford Grice et al. 2003).

4.2.1.1 cCDP in the Presence of Differential Right Censoring

A complicating factor for defining cCDP is the presence of differential right censoring in the individual components, i.e., time to onset of CDP, CIT25FWT and CI9-HPT. In such circumstances, the following algorithm will be used to define cCDP:

1. If the onset times of the individual components are all observed, then the onset time of cCDP is the earliest of the observed onset times of the individual components,

2. Otherwise if all onset times of the individual components are right censored, then the onset time of cCDP12 is right censored at the earliest of the right censored onset times of the individual components,
3. Otherwise if the earliest of the observed onset times occurs prior *or equal* to the earliest of the right censored onset times of the individual components, then the onset time of cCDP is the earliest of the observed onset times of the individual components,
4. Otherwise if the earliest of the observed onset times occurs after the earliest of the right censored onset times of the individual components, then the onset time of cCDP is set to the earliest of the observed onset times of the individual components.

In scenario 4, the approach to specify the onset time of cCDP as the earliest of the observed onset times of the individual components can in some cases overestimate the true onset time. Nonetheless, the proposed approach is pragmatic and reasonable because it leads to an estimate of the proportion of cCDP events by the end of the study that is at least as large as the observed proportion, and it has a negligible impact on treatment effect estimation if the independent censoring assumption is plausible for the individual components.

4.2.2 Main Analytical Approach for Primary Estimand

The primary estimand is the difference in time to onset of cCDP12, as expressed by the HR, between the higher and approved dose of ocrelizumab arms in patients with RMS and regardless of adherence to the randomized treatment or use of an approved MS disease-modifying therapy (DMT) medication. More specifically, the estimand has the following attributes:

Population: Patients with RMS.

Endpoint: Time to onset of cCDP12.

Summary Measure and Primary Estimator: The HR will be used to quantify the difference in time to onset of cCDP12 between study treatments, and will be estimated from a Cox proportional hazards regression model, stratified by the randomization stratification factors. The statistical significance of the estimated HR will be calculated by using the log-rank test, stratified by the randomization stratification factors.

In order to estimate the primary estimand, the following **intercurrent event** handling strategies will be applied:

Early discontinuation from study treatment: All data pre and post treatment discontinuation will be included in the analysis following the treatment-policy strategy.

Initiation of an approved MS DMT medication: All data pre and post initiation of alternative MS treatment will be included in the analysis following the treatment-policy strategy.

Pregnancy: All data pre and post pregnancy onset will be included in the analysis following the treatment-policy strategy.

Death: This is anticipated to be very rare, but in case of occurrence, it will be handled as a progression event in line with the composite strategy.

Every effort will be made to collect the data post early discontinuation from study treatment or initiation of an approved MS DMT medication until primary analysis.

Intermediate missing assessments at scheduled visits prior to the last EDSS, 9-HPT, or T25FWT assessments will not be imputed. For patients with initial disability progression based on EDSS or ≥20% increase in T25FWT or 9-HPT, data collected at the next scheduled visit or treatment discontinuation visit will be used to confirm disability progression.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP breach will not be imputed.

Patients without any prior progression events during the DBT phase will be censored at their last EDSS, 9-HPT, or T25FWT assessments (administrative censoring).

4.2.3 Sensitivity Analysis

The following sensitivity analyses of the primary estimand will be conducted according to the analysis described in Section 4.2.2.

1. The primary analysis is performed under the assumption that patients who have initial progression at their last EDSS, T25FWT or 9-HPT assessment prior to withdrawal from study are similar to those who remain in the study in terms of their risk of cCDP12. This is because such patients are censored at their last EDSS, T25FWT or 9-HPT assessment. To assess the potential violations of this assumption, the robustness of the estimated treatment effect to violations of this assumption will be investigated by repeating the primary analysis with the following imputation rules:

   For patients who had initial progression at the time of premature withdrawal from the study, multiple imputation will be used to either impute a progression event or to censor the time to onset of confirmed progression at their last EDSS, T25FWT or 9-HPT assessment prior to withdrawal from the study. The presence or absence of a progression event at the last EDSS, 9-HPT, or T25FWT assessment prior to withdrawal from the study will be imputed from a Bernoulli distribution with success probability ranging from 0-1 in increments of 0.33. This procedure will be applied within arms followed by treatment effect estimation in 100 imputed data sets. An overall estimate of the treatment effect and associated standard error for computing 95% Wald intervals will be obtained by applying Rubin's rules on the log HR scale and then back transforming to the original scale:

   $$log(HR) = 1/100 \sum_{m=1}^{100} \log(HR_m) \text{ and } SE(log(HR)) = \sqrt{V_W + \left(1 + \frac{1}{100}\right) V_B},$$

   where HR_m is the HR estimated in the mth imputed data set, V_w is the sample mean of the estimated variance of $log(HR_m)$, and V_B is the sample variance of $\log(HR_m)$.

The scenarios that result in non-significant treatment effect estimates will be identified for a tipping point analysis.

2. The primary analysis is performed under the independent censoring assumption. However, this assumption may not be plausible if treatment discontinuation affects the chance of prematurely withdrawing from study as well as cCDP12. If at least 10% participants have withdrawn from the study without having the cCDP12 event, it will be attempted to predict onset of cCDP12 event times for patients who prematurely withdraw from study based on observed post treatment discontinuation data. This prediction is deemed to be feasible only if a sufficient number of patients (i.e., at least 5% of patients) have at least one EDSS/T25FWT/9HPT assessment after treatment discontinuation. This approach of imputing missing data is aligned with the treatment policy approach for handling premature discontinuation of study treatment, since it accounts for the fact that patients who have withdrawn from study will no longer be on study treatment.

The proposed sensitivity analysis is based on the multiple imputation approach in Hartley et al. 2022, which build on the sensitivity analysis strategy in Jackson et al. 2014. Specifically, the following steps will be applied:

a) Fit a time-varying Weibull model with hazard function $\exp(\alpha)time^{\exp(\alpha)-1}\exp\{\beta_0+\beta_1X_i+\beta_2T_i+(\gamma_0+\gamma_1T_i)1(time>D_i)\}$ to time to onset of cCDP12. Here, α, β_0, β_1, β_2, γ_0, γ_1 are unconstrained parameters to be estimated, $time$ is time from randomization, X_i is a vector containing the baseline EDSS, T25FWT and 9-HPT values for patient i, T_i is an indicator such that T_i=1 if patient i was randomized to the higher dose of ocrelizumab arm and T_i=0 otherwise, and D_i is the time from randomization to discontinuation from study treatment.

b) Simulate 100 sets of $\{\alpha, \beta_0, \beta_1, \beta_2, \gamma_0, \gamma_1\}$, from the estimated asymptotic normal distribution of the maximum likelihood estimator of these parameters.

c) Let W_i and F_i be the times from randomization for patient i at the time of withdrawal from the study and end of the study. For each set of simulated parameter values, generate onset of cCDP12 event times for patients who have not had cCDP12 up to and including W_i, and where $W_i < F_i$. This will result in 100 complete analysis data sets, i.e., where cCDP12 events times are either available or are censored at F_i. In order to simulate cCDP12 events time O_i for patient i, the following formula will be used:

$$O_i = \{W_i^{\exp(a)} - \log(U_i)\exp(-\beta_0-\gamma_0-\beta_1X_i-\beta_2T_i-\gamma_1T_i)\}^{1/\exp(a)}$$

Here, U is generated from a standard uniform distribution. If $O_i > F_i$, the onset time of cCDP12 is recorded as censored at F_i in the analysis data set, otherwise it is set to O_i.

d) For each of the 100 complete analysis data sets, estimate the log HR and its associated standard error with the Cox model used for the primary analysis. Then, use Rubin's rules to combine the estimates from the 100 complete data sets, as described in sensitivity analysis 1, to obtain an overall estimate of the treatment effect on the HR scale and associated 95% Wald interval.

It is worth noting that the proposed sensitivity analysis slightly differs from that in Hartley et al. 2022. In particular, the proposed analysis follows one of the suggestions of Jackson et al. 2014, which is to use a parametric imputation model and account for the uncertainty in the estimation of the parameters indexing the imputation model by simulating from the estimated maximum likelihood estimator of these parameters. In contrast, the approach in Hartley et al. 2022 requires fitting a semi-parametric imputation model to multiple bootstrap samples in order to account for the uncertainty in the estimation of the baseline hazard and other parameters indexing the imputation model. Compared to the approach in Hartley et al. 2022, the proposed approach should therefore be less susceptible to convergence issues when there a few observed cCDP12 events after treatment discontinuation, since the imputation model only needs to be fitted to the observed data instead of multiple bootstrap samples.

In the absence of sufficient data to perform the multiple imputation described above (i.e., at least 5% of patients should have at least one EDSS/T25FWT/9HPT assessment after treatment discontinuation), the robustness of the independent censoring assumption will be evaluated using a tipping point analysis.

Multiple imputation will be used to either impute a progression event or to censor the time to onset of confirmed progression at their last EDSS, T25FWT or 9 - HPT assessment prior to withdrawal from the study. The presence or absence of a progression event at the last EDSS, 9-HPT, or T25FWT assessment prior to withdrawal from the study will be imputed from a Bernoulli distribution with success probability ranging from 0-1 in increments of 0.33. This procedure will be applied within arms followed by treatment effect estimation in 100 imputed data sets. An overall estimate of the treatment effect and associated standard error will be estimated as described in sensitivity analysis 1. The scenarios that result in non-significant treatment effect estimates will be identified for a tipping point analysis.

3. The primary analysis will be conducted using actual study days to estimate cCDP12 event times and treatment effects. Due to the periodic nature of assessments, the exact timing of cCDP12 events is unknown, as events are recorded at the next scheduled assessment after the event, resulting in interval-censored data. To assess the potential impact of using actual event days on estimation and hypothesis testing, a sensitivity analysis using interval censoring methods will be performed.

For each patient, the left and the right boundaries of the interval will be derived based on the following rules:

Situations	Left Boundary	Right Boundary
Patients who had cCDP12	The date of the last assessment that showed a disease progression-free* status	The date of the first assessment that showed cCDP12
Patients who did not have cCDP12	The date of the last assessment that showed a disease progression-free* status	Not applicable (Missing)
Patients who died and did not have cCDP12 prior to death	The date of the last assessment that showed a disease progression-free* status	Death date

cCDP12 = 12-week composite confirmed disability progression

* For patients who did not have any post-baseline assessment with disease progression-free status, the left boundary is the date of randomization.

The survival curves will be estimated using the non-parametric maximum likelihood estimate (NPMLE, Turnbull 1976) for each treatment arm.

Hypothesis testing will be performed based on the stratified log-rank test proposed by Sun (Sun 1996) to compare the cCDP12 in the treatment arms. The treatment effect will be estimated using a stratified proportional hazard regression model (Finkelstein 1986) with a parametric assumption of piecewise exponential distribution for the baseline hazard function (Friedman 1982, Royston and Parmar 2002).

The analyses listed above will be performed using the procedures PROC ICLIEFTEST and PROC ICPHREG in Statistical Analysis System (SAS) version 9.4 or a later version.

4.2.4 Supplementary Analyses

The following supplementary analyses of the primary endpoint will be conducted according to the analysis described in Section 4.2.2:

1. The primary analysis will be repeated after excluding patients who meet the following exclusion criteria/do not meet the following inclusion criteria at baseline:
    - History of primary progressive MS at screening
    - Known presence of other neurological disorders that may mimic MS
    - Any concomitant disease that may require chronic treatment with systemic corticosteroids or immunosuppressants during the course of the study
    - History of or currently active primary or secondary (non-drug-related) immunodeficiency
    - Pregnant or breastfeeding

- Previous use of anti-CD20s (including ocrelizumab), unless the last infusion was more than 2 years before screening, B-cell count is normal, and the stop of the treatment was not motivated by safety reasons or lack of efficacy
- Treatment with any investigational agent within 24 weeks prior to screening (Visit 1) or five half-lives of the investigational drug (whichever is longer), or treatment with any experimental procedure for MS (e.g., treatment for chronic cerebrospinal venous insufficiency)
- Any previous treatment with mitoxantrone, cladribine, atacicept, alemtuzumab, and daclizumab
- Previous treatment with fingolimod, siponimod, or ozanimod within 6 weeks of baseline
- Previous treatment with natalizumab within 4.5 months of baseline
- Previous treatment with interferons beta (1a or 1b), or glatiramer acetate within 2 weeks of baseline
- Previous treatment with any other immunomodulatory or immunosuppressive medication not already listed above without appropriate washout as described in the applicable local label (washout to be completed prior to baseline). If the washout requirements are not described in the applicable local label, then the wash out period must be five times the half-life of the medication. The pharmacodynamic (PD) effects of the previous medication must also be considered when determining the required time for washout
- Any previous treatment with bone marrow transplantation and hematopoietic stem cell transplantation
- Any previous history of transplantation or anti-rejection therapy
- Treatment with IV Ig or plasmapheresis within 12 weeks prior to randomization
- No Diagnosis of RMS (i.e., relapsing-remitting multiple sclerosis or active secondary progressive multiple sclerosis where patients still experience relapses) in accordance with the revised McDonald Criteria 2017
- Less than two documented clinical relapses within the last 2 years prior to screening, or no clinical relapse in the year prior to screening (with no relapse 30 days prior to screening and at baseline)
- Not neurologically stable within 30 days prior to randomization and baseline assessments
- EDSS scores at screening and baseline greater than 5.5
- Average T25FWT score over two trials at screening and over two trials at baseline respectively, above 150 seconds
- Average 9-HPT score over four trials at screening and over four trials at baseline respectively, above 250 seconds

- No documented MRI of brain with abnormalities consistent with MS at screening
- Randomized to the higher dose arm and with an incorrect classification of body weight into the strata <75 kg or ≥ 75 kg

In addition, the hypothetical strategy will be used to handle the following intercurrent events:

- Early discontinuation from study treatment
- Initiation of approved MS DMT medication
- Two consecutive missed assessments of EDSS, T25FWT and9-HPT
- Erroneously received a quantity of ocrelizumab (including no ocrelizumab) other than the group to which the patient was randomized
- Received study drug that had been mishandled (e.g., incorrect storage temperature) and was not approved for subsequent use
- Received ocrelizumab / ocrelizumab placebo but was not randomized

Under this strategy, the data after these intercurrent events will be treated as missing and the time to onset of cCDP12 will be censored at the last assessments prior to the intercurrent event times. However, the observed data after the intercurrent events will still contribute to the confirmation period for determining if an initial progression can be confirmed. This analysis estimates the treatment effect of the higher dose of ocrelizumab in a population who always receives the study treatment as planned and without consecutive missing assessments of the primary endpoint.

2. The primary analysis will be repeated after removing the individual components from cCDP12 one at a time. This analysis will determine the extent to which the treatment effect is driven solely by each of the individual components.
3. The primary analysis will be repeated after increasing the threshold for disability progression based on the performance tests from $\geq 20\%$ to $\geq 25\%$ and $\geq 30\%$. This analysis will assess the sensitiveness of the treatment effect to the influence of the performance tests on cCDP.
4. The primary analysis will be repeated after expanding the definition of the individual component endpoints in cCDP to include the following requirements between the initial progression and confirmation visit:
   - $<20\%$ decrease in T25FWT and/or 9-HPT from baseline and
   - <1.0 decrease in EDSS from baseline if baseline EDSS ≤ 5.5 or <0.5 decrease in EDSS if baseline EDSS >5.5

   This analysis will assess the sensitiveness of the treatment effect to the presence of cCDP12 events where disability improvement occurs in one or

more of the component endpoints during the confirmation period. This analysis will also be repeated for cCDP24.

5. The primary analysis will be applied to the time to onset of cCDP12 or study drug discontinuation due to AE. Thus, this supplementary analysis will incorporate discontinuation of the study drug due to AE into the definition of treatment benefit.
6. The primary analysis will be repeated (i) after censoring all patients from Russia and/or Ukraine at the onset of the Russia-Ukraine war (24th Feb 2022) and (ii) in patients not from Russia and/or Ukraine. These analyses will assess the impact of the Russia-Ukraine war on the primary analysis.

4.2.4.1 Subgroup Analyses for Primary Estimand

The following sub-group analyses will be performed via change in the population attribute of the primary estimand:

- Age at screening ($\leq$ 45 vs. > 45 years)
- EDSS (< 4.0 vs. ≥4.0)
- Region (United States vs. ROW)
- Weight at Baseline (< 75 vs. ≥75 kg)
- Presence or absence of T1Gd + lesions at baseline
- Duration since MS symptoms onset at baseline (< 3 years, ≥ 3 to ≤ 10 years, > 10 years)
- BMI at baseline (< 30 vs. $\geq$ 30 kg/m^2)
- Prior DMT at baseline (yes vs. no)
- Duration since MS onset and disability level at baseline (< 3 years & EDSS <2, < 3 years & EDSS$\geq$2, $\geq$3 years & EDSS <2, $\geq$3 years and EDSS $\geq$2)

All results will be summarized in a forest plot.

4.3 SECONDARY ESTIMANDS AND ANALYSES

Hypothesis tests for the secondary estimands as well as the hierarchical order in which they will be tested are described in Section 2.1. The p-value is interpreted as confirmatory if the primary estimand and each preceding estimand have reached the significance level of 0.05, otherwise it is reported as non-confirmatory.

All attributes for the secondary estimands are provided in Table 1. Further details on the estimands attributes and estimation methods (i.e., estimators) are provided below.

4.3.1 Estimands for Time to Event Endpoints

All secondary estimands which focus on the analysis of time to event endpoints have the same estimand attributes (except for Endpoint attribute) and estimation method as for the primary estimand (Section 4.2.2). Details on the definition of the secondary time to

event endpoints are provided in Table 1 and Section 4.2.1. Further details on the definition of the secondary time to event endpoints are provided below.

The following time to event secondary endpoints will be analyzed:

- Time to onset of cCDP24
- Time to onset of CIT25FWT12
- Time to onset of 12-week composite confirmed disability progression independent of protocol-defined relapses (cPIRA12)
- Time to onset of CDP12
- Time to onset of cCDP48
- Time to onset of CWSDMT12
- Time to onset of CIMSWS24

The time to onset of CIT25FWT12, cPIRA12, CDP12, CWSDMT12 will be defined over a 12-week confirmation window (≥84 days) for the disability progression. The conditions for the confirmation of the disability progression described for the primary estimand in Section 4.2.1 will also be applied to these endpoints.

For the disability progression the time to onset of cCDP24 and CIMSWS24 will be defined over a 24-week confirmation window (≥161 days) and the time to onset of cCDP48 will be defined over a 48-week confirmation window (≥329 days).

Time to onset of the individual components of cCDP12 (CDP12, CIT25FWT12 and CI9-HPT12) independent of protocol defined relapses is defined similarly to time to onset of the individual components of cCDP12, except that two additional conditions must be satisfied:

- The baseline reference assessment (EDSS, T25FWT, or 9-HPT values) is re-baselined at the first assessment ≥30 days after the onset of each protocol-defined relapse. Furthermore, the re-baselined disability assessment is set equal to the maximum of the disability status at the new re-baselined assessment and the previous re-baselined assessment.
- No protocol-defined relapse should occur between baseline reference assessments and within 30 days after initial progression, and 30 days prior to and after the confirmation visit.

The definitions of disability progression independent of relapses and relapse associated worsening are cited in Kappos et al. (2020).

The SDMT is a brief and easy to administer performance test that involves a simple substitution task. Using a reference key, the examinee has 90 seconds to pair specific numbers with given geometric figures. In this study, responses will be collected orally.

The 12-Item Multiple Sclerosis Walking Scale (MSWS-12) is a self-report measure of the impact of MS on the individual's walking ability.

4.3.2 Estimand for Annual Rate of Percent Change from Baseline in Total Brain Volume

Estimand: The estimand is the mean difference in annual rate of percent change from baseline in total brain volume between the higher and approved dose of ocrelizumab arms in patients with RMS where no treatment discontinuation nor initiation of alternative MS treatment can occur, according to the following attributes:

Estimator: A random coefficient regression model (RCRM) will be used to quantify mean difference in annual rate of percent change from baseline in total brain volume between the higher and approved dose of ocrelizumab arms. The mean structure of the model is as follows:

$$E\left[Y_{ij}\middle|time_{ij}, T_i, X_i\right] = time_{ij}(\beta_0 + \alpha X_i + \beta_1 T_i + u_i),$$

where Y_{ij} represents percentage change in total brain volume from baseline at visit j for patient i, $time_{ij}$ is the time from baseline in years at visit j for patient i, T_i is an indicator such that $T_i = 1$ if patient i was randomized to the higher dose of ocrelizumab arm and $T_i = 0$ otherwise, X_i is a vector containing the cube root of the baseline total brain volume and randomization stratification factors for patient i, α is a vector of regression coefficients for X_i, β_0 is the adjusted mean of the annualized rate of percentage change in total brain volume for the approved dose of ocrelizumab arm, β_1 is the adjusted mean difference in the annualized rate of percentage change in total brain volume between the higher and approved dose of ocrelizumab arms, and u_i is a subject-specific random effect that follows a mean zero normal distribution. The statistical significance of the estimated mean difference will be calculated by using the Wald test, where the standard error is estimated using the robust sandwich variance estimator.

In order to estimate the aforementioned estimand, the following **intercurrent event** handling strategies will be applied:

Early discontinuation from study treatment: All data after treatment discontinuation will be treated as missing following the hypothetical strategy.

Initiation of an approved MS DMT medication: All data after treatment discontinuation will be treated as missing following the hypothetical strategy.

Pregnancy: All data pre and post pregnancy onset will be included in the analysis following treatment-policy strategy.

Death: This is anticipated to be very rare, but in case of occurrence, it will be handled with the hypothetical strategy. That is, data after death will be treated as missing.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP breach will not be imputed.

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

```
PROC MIXED DATA=dataset empirical;
CLASS USUBJID ARMCD StratificationFactors;
MODEL AVAL_N = ARMCD Baseline StratificationFactors / solution cl;
RANDOM intercept / subject=USUBJID type=un;
run;
```

where AVAL_N is the percentage change in brain volume divided by time at each scheduled visit, and Baseline is the cube root of total brain volume at baseline.

4.3.3 Estimand for Ratio to Baseline in NfL at Week 48 for Patients Assigned to OCR HD and Approved Dose Arms

Estimand: The estimand is the mean change from baseline in NfL at Week 48 for RMS patients where no treatment discontinuation nor initiation of alternative MS treatment can occur, and within the OCR HD and approved dose arms, respectively.

Estimator: A Mixed model for repeated measures (MMRM) will be used to estimate mean change from baseline in NfL at Week 48 in the higher and approved dose of ocrelizumab arms. The MMRM will be fitted to each treatment arm separately. The dependent variable will be log base 10 of NfL, and the fixed effects in the model will include independent variables of visit (categorical variable with the following categories: Week 12, Week 24 and Week 48), log base 10 of NfL at baseline, the interaction between visit and log base 10 of NfL at baseline, and the randomization stratification factors. The model will not contain an intercept term; thus the regression coefficients of the visit levels will be an estimate of the mean change in log base 10 of NfL from baseline at the visit levels conditional on the baseline covariates. The baseline covariates in the model will also be centered, i.e., the mean of the variable over all patients in the study will be subtracted from the variable. Under this model parametrization, the estimate of the mean change from baseline at week 48 will be computed as 10^ (regression coefficient of Week 48), which is based on an estimate of the geometric mean.

An unstructured variance−covariance structure will be applied to model the within-patient errors. The restricted maximum likelihood (REML) method will be used for estimation of variance components. The statistical significance of the estimated mean change from baseline will be calculated by using the Wald test.

In order to estimate the aforementioned estimand, the following **intercurrent event** handling strategies will be applied:

Early discontinuation from study treatment: All data after treatment discontinuation will be treated as missing following the hypothetical strategy.

Initiation of an approved MS DMT medication: All data after treatment discontinuation will be treated as missing following the hypothetical strategy.

Pregnancy: All data pre and post pregnancy onset will be included in the analysis following treatment-policy strategy.

Death: This is anticipated to be very rare, but in case of occurrence, it will be handled with the hypothetical strategy. That is, data after death will be treated as missing.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP breach will not be imputed.

4.4 EXPLORATORY ESTIMANDS AND ANALYSIS

This section defines estimands for exploratory efficacy endpoints listed in Table 2.

4.4.1 Estimands for Time-to-Event Endpoints

Estimand: The estimand is the difference in endpoint, as expressed by the HR, between the higher and approved dose of ocrelizumab arms in patients with RMS and regardless of adherence to the randomized treatment or use of an approved MS DMT medication.

Endpoints:

- Time to onset of 24-week confirmed ≥20% increase in T25FWT (CIT25FWT24)
- Time to onset of 24-Week confirmed disability progression in EDSS (CDP24)
- Time to onset of 12-week confirmed ≥20% Increase in 9-HPT (CI9-HPT12)
- Time to onset of 24-week confirmed ≥20% Increase in 9-HPT (CI9-HPT24)
- Time to onset of 12-week confirmed ≥ 8-point worsening in SDMT
- Time to onset of cCDP12 or 12-weekconfirmed ≥ 4-point worsening in SDMT
- Time to onset of cCDP12 or 12-week confirmed ≥ 8-point worsening in SDMT
- Time to onset of cCDP12 individual components independent of protocol-defined relapses
- Time to onset of cCDP24 and cCDP24 individual components independent of protocol-defined relapses
- Time to onset of 12-week confirmed protocol-defined relapse associated worsening in cCDP12 and individual components

Time to onset of protocol-defined relapse associated worsening in the individual components of cCDP12 is defined similarly to time to onset of the individual components of cCDP12, except that one additional condition must be satisfied:

- Progression must occur within 90 days of a protocol defined relapse.

If progression occurs and it does not satisfy the above condition, then 1) the progression is not considered a protocol-defined relapse associated worsening event, 2) time to

onset of protocol-defined relapse associated worsening is censored at the time of the progression.

Time to onset of protocol-defined relapse associated worsening in cCDP12 is defined as the time to onset of the first occurrence of protocol-defined relapse associated worsening in the individual components of cCDP12.

Summary Measure and Primary Estimator: The HR will be the summary measure and the Cox proportional hazards model will be the primary estimator. The estimation method will be the same as for the primary estimand (Section 4.2.2).

In order to estimate the aforementioned estimand, the same **intercurrent event** handling strategies as those for the primary endpoint in Section 4.2.2 will be applied.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP breach will not be imputed (including following death).

4.4.2 Estimands for Continuous Endpoints at Scheduled Visits

Estimand: The estimand is the mean difference in the endpoint between the higher and approved dose of ocrelizumab arms in patients with RMS and regardless of adherence to the randomized treatment or use of an approved MS DMT medication.

Endpoints:

- Change from baseline in EDSS score at Week 48 and 120
- Change from baseline in Multiple Sclerosis Impact Scale-29 (MSIS-29) physical scale at Week 48 and 120
- Change from baseline in MSIS-29 psychological scale at Week 48 and 120
- Change from baseline in Quality of Life in Neurological Disorders (Neuro-QOL) Upper Extremity Function Form at Week 48 and 120
- Change from baseline in Modified Fatigue Impact Scale at Week 48 and 120

Summary Measure and Primary Estimator: MMRM will be used to quantify mean difference in the endpoints between the higher and approved dose of ocrelizumab arms. 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 outcome, treatment-by-visit interaction, baseline outcome by visit interaction, along with the randomization stratification factors. An unstructured variance-covariance structure will be applied to model the within-patient errors. The restricted maximum likelihood (REML) method will be used for estimation of variance components. The statistical significance of the estimated mean difference will be calculated by using the Wald test.

In order to estimate the estimand of interest, the following **intercurrent event** handling strategies will be applied:

Early discontinuation from study treatment: All data available will be included in the analysis following the treatment-policy strategy.

Initiation of an approved MS DMT medication: All data available will be included in the analysis following the treatment-policy strategy.

Pregnancy: All data pre and post pregnancy onset will be included in the analysis.

Death: This is anticipated to be very rare, but in case of occurrence, it will be handled with the composite strategy. That is, data after death will be treated as the worst possible score.

Only data up until Week 120 (inclusive) will be used in the analysis.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP will not be imputed (including following death).

For MSIS-29, if there are missing individual items and the respondents have a minimum of 10 completed items in the physical scale, or a minimum of 5 completed items in the psychological scale, the following imputation process will be applied to handle the missing individual items. Missing scores in the specific scale are imputed as the average of the scores of all (completed) items in that scale. For example, if a patient has completed 15 items (item 1-15) in the physical scale and the values for item 16-20 are missing, then first sum the completed items 1-15 and then divide by 15 to get their respondent–specific mean score. Then use this value as the score for each of the missing 5 items 16-20. Generate a total score as usual by summing the values of the 15 completed items 1-15 and the 5 imputed items 16-20 based on the score calculation for physical scale (above).

4.4.3 Estimands for Rate of Change of Continuous Endpoints

Estimand: The estimand is the mean difference in the endpoint between the higher and approved dose of ocrelizumab arms in patients with RMS where no treatment discontinuation nor initiation of alternative MS treatment can occur.

Endpoints:

- Annual rate of percentage change in thalamic volume
- Annual rate of change from baseline in total radius of T1-hypo-intense lesions (black holes)
- Annual rate of percentage change in volume of spinal cord (upper part of the spine)

The same population definition, intercurrent events handling strategies and estimator will apply as for the annual rate of percent change from baseline in total brain volume (Section 4.3.2).

For annual rate of change from baseline in total radius of T1-hypo-intense lesions (black holes), Y_{ij} represents change in the cube root of total volume of T1-hypo-intense lesions from baseline at visit j for patient i.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP breach will not be imputed (including following death).

4.4.4 Estimands for Proportion of Responders at Scheduled Visits

Estimand: The estimand is the difference in the endpoint, as expressed by the odds ratio of being a responder, between the higher and approved dose of ocrelizumab arms in patients with RMS and regardless of adherence to the randomized treatment or use of an approved MS DMT medication.

Endpoints:

- Proportion of patients with no change in overall severity, improvement or worsening measured with patient global impression of severity at Week 48 and 120
- Proportion of patients with no change in overall MS symptoms, improvement or worsening measured with patient global impression of change at Week 48 and 120
- Proportion of patients with no change in upper limb function, improvement or worsening measured with patient global impression of change of upper limb function at Week 48 and 120

Summary Measure and Primary Estimator: Ordered logistic regression will be applied to the various responder definitions at each scheduled visit between the higher and approved dose of ocrelizumab arms. The fixed effects in the model will include independent variables of randomized treatment, baseline endpoint value, and the randomization stratification factors. The statistical significance of the estimated odds ratio will be calculated with the Wald test.

In order to estimate the estimand of interest, the same **intercurrent event** handling strategies as for primary estimand in Section 4.2.2 will be applied.

Only data up until Week 120 (inclusive) will be used in the analysis.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP breach will not be imputed (including following death).

4.4.5 Estimands for Expected Counts at each Scheduled Visit

Estimand: The estimand is the ratio of the mean endpoint between the higher and approved dose of ocrelizumab arms in patients with RMS and in the absence of treatment discontinuation and initiation of alternative MS treatments.

Endpoints:

- Total number of new or enlarging T2 lesions per MRI scan over the 120-week treatment period and at each scheduled visit
- Total number of T1Gd+ lesions over the 120-week treatment period and at each scheduled visit
- Total number of new T1-hypo-intense lesions (black holes) over the 120-week treatment period and at each scheduled visit

Summary Measure and Primary Estimator: Separate negative binomial models will be fitted to the data on the endpoints at each scheduled visit and for the total count over the 120-week treatment period. The mean function of the model will be parameterized using the log link function and the linear predictor will include as an offset either the log number of MRI scans for T1Gd+ lesions or log time in years of scan relative to baseline for new or enlarging T2 lesions and T1-hypo-intense lesions, an indicator variable for the assigned treatment and the randomization stratification factors as independent variables. The statistical significance of the estimated ratio of means will be calculated by using the Wald test.

In order to estimate the aforementioned estimand, the same **intercurrent event** handling strategies as those for brain volume in Section 4.3.2 will be applied.

Only data up until Week 120 (inclusive) will be used in the analysis.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious GCP breach will not be imputed (including following death).

4.4.6 Estimand for Annualized Relapse Rate

Estimand: The estimand is the difference in the ratios of the annualized relapse rate, expressed by the rate ratio, between the higher and approved dose of ocrelizumab arms in patients with RMS and regardless of adherence to the randomized treatment or use of an approved MS DMT medication.

Summary Measure and Primary Estimator: The rate ratio will be used to quantify the ratio of the annualized relapse rate, and will be estimated from fitting a negative binomial regression model to the number of relapses. The mean function of the model will be parameterized by using the log link function and the linear predictor will include log time in study (years) as an offset, an indicator variable of the randomized treatment and the randomization stratification factors as independent variables. The statistical significance of the estimated rate ratio will be calculated by using the Wald test.

In order to estimate the estimand of interest, the following **intercurrent event** handling strategies will be applied:

Early discontinuation from study treatment: All available data pre and post treatment discontinuation will be included in the analysis following the treatment-policy strategy.

Initiation of an approved MS DMT medication: All data pre and post initiation of alternative MS treatment will be included in the analysis following the treatment-policy strategy.

Pregnancy: All data pre and post pregnancy onset will be included in the analysis following the treatment-policy strategy.

Death: This is anticipated to be very rare, but in case of occurrence, it will be handled with the hypothetical strategy. That is, data on relapse after death will be treated as missing.

Missing data resulting from patients that withdraw their consent to participate in the study and/or a serious CGP breach will not be imputed.

4.5 SAFETY ANALYSES

The safety analysis set (SAS) is described in Section 3. All safety parameters will be summarized and presented in tables on the basis of this analysis set. Randomized patients who receive incorrect therapy will be summarized in the group according to the rules laid out in Section 3. Patients who received at least one dose of ocrelizumab above 600 mg will be grouped in the higher dose of ocrelizumab arm.

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

All patients who received at least one infusion (partial or complete) of study drug with patients grouped according to actual treatment received will be included in the analysis.

- The safety data will be summarized regardless of premature study treatment discontinuation or switch to an approved MS DMT medication (including commercial ocrelizumab).
- Additionally, the safety data may also be summarized regardless of premature study treatment discontinuation but until switching to an approved MS DMT medication (including commercial ocrelizumab).

Safety will be assessed through summaries of exposure to study treatment, AEs, changes in laboratory test results, and changes in vital signs. All verbatim AE terms will be mapped to MedDRA thesaurus terms, and AE severity will be graded according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v5.0. All AEs, SAEs, AEs leading to death, AEs of special interest, and AEs leading to study treatment discontinuation that occur on or after the first dose of study treatment (i.e., treatment-emergent AEs) will be summarized by mapped term,

appropriate thesaurus level, and severity grade. For events of varying severity, the highest grade will be used in the summaries. Deaths and cause of death will be summarized. Relevant laboratory and vital sign (pulse rate, and blood pressure) data will be displayed by time, with grades identified where appropriate. Additionally, a shift table of selected laboratory tests will be used to summarize the baseline and maximum post baseline severity grade. Changes in vital signs will be summarized.

4.5.1 Extent of Exposure

Study treatment exposure such as treatment duration, total dose received, and number of doses will be summarized with descriptive statistics.

4.5.1.1 Details in Defining the Number of Doses and Total Dose Received

The first dose of ocrelizumab is given as two infusions administered 2 weeks apart. If a patient receives any infusion at Dose 1, the patient is counted as having received the first dose.

If a dose is completely missed instead of delayed, the next dose number will be consecutive to the previous dose.

The total cumulative dose is capped at 100%, so the actual dosage for each infusion does not exceed 600 mg or 1200/1800 mg, respectively. The first dose is given in two infusions and the cumulative dose for each of these infusions is also capped at 100%.

4.5.1.2 Treatment Duration and Duration under Observation Definitions

Treatment duration will be calculated as follows:

(Date of the last recorded treatment observation during the DBT phase* − Date of first dose of the DBT phase), where* is the Earliest 1) date of last dose received during the DBT phase + 24 weeks; 2) date of CCOD for the primary analysis reporting; 3) date of treatment discontinuation as described in Section 4.1 for patients who discontinued study treatment; 4) date of subject withdrawal from the study; 5) date of death. The duration under observation for a patient will be calculated as follows:

(The latest recorded date in all eCRF and non eCRF prior to CCOD − Date of first infusion in the first dose) + 1

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 nth 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.

4.5.2 Adverse Events

AEs will be defined as all AEs including IRRs and serious MS relapses, but excluding non-serious MS relapses. Therefore, those AEs recorded on the AE and IRR case report form (CRF) pages will be included. Refer to Section 4.5.7 for further guidance on MS relapses considered serious AEs.

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 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 SAS, 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 or with CRF flag "Event occurred prior to first study drug administration" equals "no". 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, non-imputed start date will be assumed to be treatment emergent unless the AE has a complete, non-imputed 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 (≥1%) will be presented by SOC and 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 remain in the DBT phase without treatment. 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 (see Section 4.5.4), opportunistic infections (OIs; see Section 4.5.5), and malignancies (see Section 4.5.6).

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

(Total number of AEs ÷Total number of patient-years) × 100

- The 95% CI for the number of AEs per 100 patient-years will be calculated using the exact method based on the Poisson distribution (Sahai and Khurshid 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 Chi-squared distribution with 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 dose and by treatment (including all doses).

To assess the impact of coronavirus disease-2019 (COVID-19) pandemic, AEs may be summarized both including and excluding COVID-19 related AEs.

4.5.3 Infusion Related Reactions

The occurrence of 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. 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.

Onset times of IRRs, if known, are recorded on the CRF. 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 (i.e., During infusion, Event occurred within 24 hours after end of IMP infusion, time not available).

Additionally, similar tables will be presented by the pre-medication subgroup (methylprednisolone plus antihistaminics, or methylprednisolone plus antihistaminics and analgesic/antipyretic).

Symptoms of IRRs and symptoms of serious IRRs will be presented by infusion. Symptoms of IRRs that led to interruption 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 and symptoms of IRR experienced within 24 hours after end of IMP infusion will also be presented.

AEs other than IRRs experienced by the patient within 24 hours of an infusion will be presented by infusion (may include the day of the infusion and 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 treatments for IRRs will be summarized.

4.5.4 Infections

Infections will be defined from the AE data using the MedDRA SOC of “Infections and infestations”. Infections will be classified according to the pathogen type (e.g., bacterial, fungal, viral, parasitical, unknown).

An infection will be defined as serious if the event is an SAE as defined in Section 4.5.2.

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 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/year of exposure 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.

4.5.5 Opportunistic Infections

Opportunistic infections (OIs) will be defined using the Standardized MedDRA query (SMQ) of “Opportunistic infections (narrow).”

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 and by dose/year of exposure.

4.5.6 Malignancies

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

4.5.7 Multiple Sclerosis Relapses

AE outputs will include all relapses recorded on the AE CRF page, and therefore will only include serious relapses.

Information related to a protocol-defined relapse will be captured on a clinical relapse event CRF page. A clinical relapse will qualify as a protocol-defined relapse if it satisfies the definition in Section 4.2.1.

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.

4.5.8 Pregnancies

Pregnancy information will be summarized in individual patient listings.

4.5.9 Laboratory Data

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 (including those for hematology, chemistry and Ig levels).

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.

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 time point, the number and percentage of patients with immunoglobulin levels lower than the lower limit of normal will be presented.

HBV DNA

Hepatitis B virus (HBV) DNA (polymerase chain reaction [PCR]) in patients who enrolled with negative HBsAg and positive total HBcAb will be listed.

4.5.10 Vital Signs

Changes from study baseline in vital signs including systolic and diastolic blood pressure, and pulse rate will be summarized by visit and group.

Changes from pre-infusion baseline to post-infusion time points will also be summarized for each infusion.

4.6 OTHER ANALYSES

4.6.1 Summaries of Conduct of Study

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

- Summary of major protocol deviations
- Summaries of FAS and SAS, including numbers of patients in each analysis set
- Summary of subject disposition, including the number of treatment doses received, reasons for withdrawal from study treatment
- Consort diagram of the study disposition including intercurrent events
- Table of the number of patients switching to other disease modifying therapies
- Figure of cumulative enrollment over time
- Summary and listing of intercurrent events
- Summary and Kaplan-Meier plots of:

    1. 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)
    2. Time to discontinuation from the study. (All other patients will be censored at their last assessment during the study)

4.6.2 Summaries of Treatment Group Comparability/Demographics and Baseline Characteristics

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 available eCRF data:

- Age (years) at screening
- Age stratification category ($\leq$45 years; $>$45 years)
- Sex (male patients, female patients)

- Race (White, Black or African American, Other)
- Ethnicity (Hispanic or Latino, non-Hispanic or Latino)
- Weight at Baseline (kg)
- Weight stratification category (<75 kg; ≥ 75 kg)
- Body mass index (measured in kg/m^2) – derived from weight and height, i.e. weight (in kilograms) divided by height (in meters) squared.
- Region stratification category (United States; ROW)
- EDSS stratification (< 4.0; ≥ 4.0)

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

Note that race cannot be collected in France.

Baseline disease characteristics:

- EDSS (continuous)
- 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)
- T25FWT (continuous)
- 9-HPT (continuous)
- SDMT score (continuous)
- MSWS-12 (continuous)

MS disease history:

- Duration since MS symptom onset (calculated in years; i.e., divide duration in days by 365.25)
- Duration since MS symptom onset category (≤ 1 years, >1 to ≤ 3 years, >3 to ≤ 5 years, >5 to ≤ 10 years, 10 to ≤ 15 years, >15 to ≤ 20 years, >20 years)
- Duration since RMS diagnosis (calculated in years up to randomization date)
- Current phenotype of MS (RRMS, aSPMS)

- Treatment with any MS disease-modifying therapy prior to the baseline visit (yes [treatment-experienced patients], no [naive patients])

  MS disease-modifying therapy includes: fingolimod, ozanimod, siponimod, mitoxantrone, dimethyl fumarate, glatiramer acetate, interferon (beta 1a; beta 1b; peginterferon beta 1a), teriflunomide, laquinimod, daclizumab (license suspended), natalizumab, ofatumumab, alemtuzumab, rituximab, ocrelizumab, belimumab, atacicept, cladribine, ublituximab, fenebrutinib, Evobrutinib, ATX-MS-1467, Diroximel fumarate, Ponesimod, Masitinib.

Baseline MRI data:

- Number of gadolinium-enhancing T1 lesions at baseline (continuous and categorical analysis [0, 1, 2, 3, $\geq$4])
- Whole-brain volume at baseline
- T2 lesion volume at baseline
- Number of T2 lesions at baseline (continuous)

Baseline Lab data:

- NfL at baseline

4.6.3 Summaries of impact of Russia-Ukraine War on the Study

The following analyses will be conducted to evaluate the impact of the Russia-Ukraine war on the study conduct:

- Subject disposition stratified by country (Russia, Ukraine and rest of the world)
  - Number and proportion of patients that missed at least two consecutive EDSS, 9HPT and T25FWT assessments
  - Mean and median time on study
  - Number and proportion of patients that completed the DBT phase
  - Number and proportion of patients that discontinued study treatment
  - Number and proportion of patients that withdrew from study
- Protocol deviations stratified by country
  - Number and proportion of patients with at least one major protocol deviation
  - Average number of major protocol deviations per patient
- Adverse and serious adverse event rates before and after the onset of the war stratified by country

- Summary statistics for laboratory parameters assessed at local and central laboratories

Appendix 1 provides details of the tables that will contain the information described above.

4.6.4 Pharmacokinetic Analyses

The pharmacokinetic analysis set (PAS) is described in Section 3. All PK parameters will be summarized and presented in tables on the basis of this analysis set.

Nonlinear mixed-effects modeling will be used to analyze the sparse concentration–time data of ocrelizumab. All measurements of 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 or if data are unavailable, not plausible, or incomplete that may interfere with PK evaluation. Excluded cases will be documented together with the reason for exclusion. 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, anti-drug antibodies (ADAs), and baseline CD19+ B cell count on these parameters will be investigated.

The mixed-effects modeling analyses will be reported separately.

4.6.5 Pharmacodynamic Analysis

The PD objective for this study is to characterize the ocrelizumab PD profile on the basis of the following endpoints:

- B-cell levels in blood (including comparing the degree of B-cell depletion between the doses)
- Proportion of patients achieving 5 or less B-cells per microliter of blood
- Proportion of patients achieving 0.4 or less B-cells per microliter of blood

The analyses will be performed for the SAS as described in Section 3 with and without patients from Russia, where there were operational challenges to obtain B-cell measurements.

4.6.6 Immunogenicity Analyses (Anti-Drug Antibodies)

The immunogenicity analysis set (IAS) set is described in Section 3. All immunogenicity parameters will be summarized and presented in tables on the basis of this analysis set.

The numbers and proportions of ADA-positive patients and ADA-negative patients at baseline (baseline prevalence) and after drug administration (post baseline incidence) will be summarized 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 time points where ADA samples were collected and analyzed will be

presented. A listing by treatment of ADA data will be presented for patients with at least one ADA sample.

When determining post-baseline incidence, patients are considered to be ADA positive if they are ADA negative or have missing data at baseline but develop an ADA response following study drug exposure (treatment-induced ADA response), or if they are ADA positive at baseline and the titer of one or more post baseline samples is at least 0.60 titer unit greater than the titer of the baseline sample (treatment-enhanced ADA response). Patients are considered to be ADA negative if they are ADA negative or have missing data at baseline and all post baseline samples are negative, or if they are ADA positive at baseline but do not have any post baseline samples with a titer that is at least 0.60 titer unit greater than the titer of the baseline sample (treatment unaffected).

The relationship between ADA status and safety, efficacy, PK, and biomarker endpoints may be analyzed and reported descriptively.

4.6.7 Biomarker Analyses

The SAS is described in Section 3. All biomarker parameters will be summarized and presented in tables on the basis of this analysis set.

Biomarkers will be assessed at baseline and subsequent time points following administration of ocrelizumab. Biomarkers will be presented as absolute value over time and/or percent change relative to baseline over time. Biomarker levels at baseline or over time may be compared with efficacy or safety measurements to assess prognostic or predictive properties. Descriptive or summary statistics will be used to describe biomarker assessments. Samples taken at screening and baseline before first ocrelizumab dose may be averaged to better normalize post-dose endpoints taking into account normal variability.

The following endpoints may be included in the biomarker analysis:

- Levels of soluble biomarkers including but not limited to NfL and/or IL-6 in blood (plasma and/or serum)
- Levels of B or T cell subsets in blood, including but not limited to CD19+ IgD, CD27, CD38, CD4, CD8, CD3, parameters to identify B or T naive, memory and/or B plasma blast/plasma cell subsets
- DNA genotype of patients to include but not be limited to FcγR3A and human leukocyte antigen genotype. Collection and submission of blood samples for whole genome sequencing or whole exome sequencing is contingent upon the review and approval of the exploratory research by each site's Institutional Review Board or Ethics Committee and, if applicable, an appropriate regulatory body

A Biomarker Analysis Plan (BAP) that describes all planned analyses of biomarker data will be prepared and finalized prior to the data base analysis lock.

4.6.8 Health Status Analyses

The EQ-5D-5L is a validated instrument to describe and measure health (EuroQol Group 1990; Brooks 1996; Herdman et al. 2011; Janssen et al. 2013). There are two components to the EQ-5D-5L: a five-item descriptive system questionnaire (EQ-5D-5L) that assesses mobility, self-care, usual activities, pain/ discomfort, and anxiety and depression and a visual analog scale (EQ VAS) that rates the overall health. Absolute and change from baseline at each scheduled visit in EQ VAS scale will be presented by treatment arm. Further exploratory analyses of EQ-5D-5L may include the calculation of EQ-5D-5L utility index score and clinical measurements that may to support pharmacoeconomic modeling. These exploratory analyses will not be included in the CSR.

4.6.9 COVID-19 Vaccine Immune Response Analyses

Frequency of patients with positive SARS-CoV-2 antibody response and positive SARS-CoV-2 T cell response following vaccination within the first 3 months following primary or booster vaccination, within the first 6 months and/or 1 year, will be presented by treatment arm.

Descriptive summary statistics and listings will be provided for the antibody titers and T cell levels.

Further details will be outlined in the BAP.

4.7 CHANGES TO PROTOCOL-PLANNED ANALYSES

Pregnancy has been added as an intercurrent event in the definition of the primary estimand in this SAP. The rationale for this change is that several pregnancies were identified after the protocol was finalized, and the presence of pregnancies can impact the interpretation of the study results.

5. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document.

6. REFERENCES

Bohannon RW. Comfortable and maximum walking speed of adults aged 20−79 years: reference values and determinants. Age Ageing 1997;26:15−19.

Brooks R, on behalf of the EuroQol Group. EuroQol: the current state of play. Health Policy. 1996;37:53–72.

D'Souza M, Yaldizli Ö, John R, et al. Neurostatus e-Scoring improves consistency of Expanded Disability Status Scale assessments: A proof of concept study. Mult Scler Houndmills Basingstoke Engl 2017:597−603.

EuroQol Group (1990). EuroQol--a new facility for the measurement of health-related quality of life. Health policy (Amsterdam, Netherlands), 16(3), 199–208

Finkelstein, D. M. (1986). A proportional hazards model for interval-censored failure time data. Biometrics, 42, 845-854.

Friedman, M. (1982). Piecewise exponential models for survival data with covariates. Annals of Statistics, 10, 101–113.

Hartley B, Drury T, Lettis S, et al. Estimation of a treatment policy estimand for time to event data using data collected post discontinuation of randomised treatment. Pharm Stat. 2022; 21:612−24.

Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research. 2011 Dec;20:1727-36.

Jackson, D., White, I. R., Seaman, S., et al. Relaxing the independent censoring assumption in the Cox proportional hazards model using multiple imputation, Statist. Med. 2014; 27;4681–94.

Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res. 2013;22(7):1717-1727.

Kappos L, Bar-Or A, Cree BAC, et al. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. Lancet 2018; 391:1263−73.

Kappos L, Wolinsky JS, Giovannoni G, et al. Contribution of Relapse-Independent Progression vs Relapse-Associated Worsening to Overall Confirmed Disability Accumulation in Typical Relapsing Multiple Sclerosis in a Pooled Analysis of 2 Randomized Clinical Trials. JAMA Neurol. 2020; 77:1132−40.

Oxford Grice KA, Vogel KA, Le V, et al. Adult Norms for a Commercially Available Nine Hole Peg Test for Finger Dexterity. Am J Occup Ther 2003; 57:570−73.

Royston, P. and Parmar, M. K. B. (2002). Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic

modelling and estimation of treatment effects. Statistics in Medicine, 21, 2175–2197.

Sahai H, Khurshid A. Confidence intervals for the mean of a Poisson distribution: A review. 1993. Biometrical J, 35: 857−67.

Sun, J. (1996). A non-parametric test for interval-censored failure time data with application to AIDS studies. Statistics in Medicine, 15, 1387-1395.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored, and truncated data. Journal of the Royal Statistical Society, Series B 38, 290–295.

Appendix 1 Tables to Display the Impact of the Ukraine-Russia War on the Study

Proposed table to assess subject disposition by country:

Table displaying subject disposition by country.

Variable	ROW (N=)	UKR (N=)	RUS (N=)
Missed assessments of EDSS, 9HPT and T25FWT			
Number of patients with at least two consecutive missed assessments (%)			
Time in study			
Mean (SD)			
Median			
Proportion of time in study (Time in study/Time from randomization date to clinical cutoff date for the primary analysis)			
Mean (SD)			
Median			
DBT completion status			
Ongoing (%)			
Discontinued (%)			
Study withdrawal reason			
War dependent			
War independent			
Treatment completion status			
Ongoing (%)			
Discontinued (%)			
Treatment discontinuation reason			
Adverse event (%)			
Pregnancy (%)			
Death (%)			
Lack of efficacy (%)			
Loss to follow-up (%)			
Protocol deviation (%)			
Withdrawal by subject (%)			
Study terminated by Sponsor (%)			
Physician decision (%)			
Other (%)			

9–HPT = 9–hole peg test; EDSS = expanded disability status scale; T25FWT = timed 25-foot walk test; DBT = double-blind treatment.

Proposed table to assess major protocol deviation status by country:

Table displaying number of patients with at least one and total number of major protocol deviations by country.

	ROW (N=)	**UKR (N=)**	**RUS (N=)**
Number of patients with at least one of the following			
Major PD (%)			
War independent major PD (%)			
War dependent major PD (%)			
Total number of the following			
Major PD (per patient)			
War independent major PD (per patient)			
War dependent major PD (per patient)			

Proposed table to assess change in adverse and serious adverse event rates by country:

Table displaying adverse and serious adverse event rates before and after the onset of the Ukraine/Russia conflict (24/02/2022)

	ROW (N=)	UKR (N=)	RUS (N=)
Adverse events per 100 patient years			
Before the start of the war (number of events/total follow-up time)	(xx/xx years)	(xx/xx years)	(xx/xx years)
After the start of the war (number of events/total follow-up time)	(xx/xx years)	(xx/xx years)	(xx/xx years)
Serious adverse events per 100 patient years			
Before the start of the war (number of events/total follow-up time)	(xx/xx years)	(xx/xx years)	(xx/xx years)
After the start of the war (number of events/total follow-up time)	(xx/xx years)	(xx/xx years)	(xx/xx years)

Table displaying summary statistics of selected lab parameters by countries and laboratory.

sVisit	ROW	UKR		RUS	
	Central Lab	Central Lab	Local Lab	Central Lab	Local Lab
Baseline N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					
Week 12 N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					
Week 24 N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					
Week 36 N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					
Week 48 N Mean (sd) Median (Q1;Q3) Min-Max					

……. N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					
All visit pre onset of the war (**24/02/2022**) N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					
All visit post onset of the war (**24/02/2022**) N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					
All visits N Mean (sd) Mean ± 2 sd Median (Q1;Q3) Min-Max					

Signature Page for Statistical Analysis Plan - BN42082 - Published
System identifier: RIM-CLIN-986000

Approval Task	[redacted] Company Signatory 18-Feb-2025 08:43:34 GMT+0000