

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

761053Orig1s000

STATISTICAL REVIEW(S)



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION ADDENDUM

NDA/BLA #:	BLA 761,053
Drug Name:	Ocrelizumab
Indication(s):	Multiple Sclerosis
Applicant:	Genentech
Date(s):	Date of Submission: April 28, 2016 PDUFA Date: December 28, 2016
Review Priority:	Priority Review
Biometrics Division:	Division of Biometrics I
Statistical Reviewer:	Sharon Yan
Concurring Reviewers:	Kun Jin, Team Leader Jim Hung, Director
Medical Division:	Neurology
Clinical Team:	Lawrence Rodichok, Clinical Reviewer Gerard Boehm, Safety Reviewer John Marler, Clinical Team Leader Eric Bastings, Deputy Director Billy Dunn, Director
Project Manager:	Nahleen Lopez

The following analysis was performed for the PPMS study 25046 for the purpose of 1) to evaluate the likelihood of confirming a disability progression once an onset has occurred, and 2) to assess the variation of outcomes from patients who withdrew from the study after an initial disability progression event before the confirmation visit. Refer to the main body of the Statistical Review and Evaluation for more details of efficacy analysis and related information.

- 1) In the PPMS Study 25046, a total of 256 confirmed disability progression (CDP) events were observed. Among them, 235 events were actually confirmed and 21 events were imputed as CDP events due to patients who had an initial onset and then withdrew before possible confirmation could be made at a second visit 12 weeks after. In addition to the 256 CDP events observed, 76 patients had an initial onset and the event was not confirmed at a 12-week assessment after the initial onset. Therefore, out of the 332 (235+21+76) patients who had an initial onset, 76 patients did not have CDP. The likelihood of having an onset not confirmed in this study is $76/332=22.89\%$.
- 2) Under the assumption that 22.89% of the 21 patients who had an initial onset event and then withdrew from the study would not likely to be confirmed, 5 patients ($21 \times 22.89\% = 4.81$ rounded to the nearest whole number) were randomly selected from these 21 patients and their event status were modified from the CDP to censoring at the withdrawal. The same primary model of the log-rank test was then applied to the modified data to obtain the p-value from the treatment comparison.

The procedure was repeated for 500 times, each with a different set of 5 patients randomly chosen to have their event status modified. The following table presents the descriptive statistics of the 500 p-values.

Descriptive Statistics of 500 p-values	
Mean	0.050
Minimum	0.0177
Maximum	0.0931
90% Range	0.0256 – 0.0701

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

XIAORONG YAN
02/10/2017

KUN JIN
02/10/2017
I concur with the review.

HSIEN MING J HUNG
02/10/2017



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:	BLA 761,053
Drug Name:	Ocrelizumab
Indication(s):	Multiple Sclerosis
Applicant:	Genentech
Date(s):	Date of Submission: April 28, 2016 PDUFA Date: December 28, 2016
Review Priority:	Priority Review
Biometrics Division:	Division I
Statistical Reviewer:	Sharon Yan
Concurring Reviewers:	Kun Jin, Team Leader Jim Hung, Director
Medical Division:	Neurology
Clinical Team:	Lawrence Rodichok, Clinical Reviewer Gerard Boehm, Safety Reviewer John Marler, Clinical Team Leader Eric Bastings, Deputy Director Billy Dunn, Director
Project Manager:	Nahleen Lopez

Table of Contents

1	EXECUTIVE SUMMARY	6
2	INTRODUCTION.....	7
2.1	OVERVIEW.....	7
	SOURCE: REVIEWER’S SUMMARY	8
2.2	DATA SOURCES	8
3	STATISTICAL EVALUATION.....	8
3.1	DATA AND ANALYSIS QUALITY	8
3.2	EVALUATION OF EFFICACY.....	8
3.2.1	<i>Evaluation of Efficacy for Studies WA21092 and WA21093</i>	<i>8</i>
3.2.1.1	Study Design	8
3.2.1.2	Study Endpoints	10
3.2.1.2.1	Definition of Efficacy Endpoints.....	10
3.2.1.3	Statistical Methodologies	12
3.2.1.3.1	General Consideration	12
3.2.1.3.2	Analyses of the Primary Endpoints	14
3.2.1.3.3	Analysis of Secondary Endpoints	14
3.2.1.4	Study Results of Protocols WA21092 and WA21093.....	17
3.2.1.4.1	Patient Disposition.....	17
3.2.1.4.2	Patient Demographics.....	18
3.2.1.4.3	Patient Baseline Disease Characteristics	18
3.2.1.5	Efficacy Results of WA21092 and WA21093 Reported by the Sponsor	19
3.2.1.5.1	Annualized Relapse Rate (ARR) – Primary Endpoint.....	19
3.2.1.5.2	Secondary Endpoints	22
3.2.1.6	Efficacy Results from Reviewer’s Analyses	25
3.2.1.6.1	Results from Analysis of Primary Endpoint ARR	25
3.2.1.6.2	Analysis of 12-Week CDP.....	25
3.2.1.6.3	Analysis of Other Secondary Endpoints.....	26
3.2.1.6.4	Summary of Reviewer’s Analysis	27
3.2.2	<i>Evaluation of Efficacy for Protocol WA25046</i>	<i>28</i>
3.2.2.1	Study Design	28
3.2.2.2	Study Endpoint.....	29
3.2.2.3	Statistical Analysis Methods	29
3.2.2.3.1	General Considerations.....	29
3.2.2.3.2	Analysis of the Primary Endpoint.....	30
3.2.2.3.3	Analysis of Secondary Endpoints	31
3.2.2.4	Patient Disposition, Demographic and Baseline Characteristics.....	33
3.2.2.4.1	Patient Disposition.....	33
3.2.2.4.2	Patient Demographics.....	34
3.2.2.4.3	Patient Baseline Disease Characteristics	35
3.2.2.5	Efficacy Results of Study WA25046 Reported by the Sponsor	36
3.2.2.5.1	Analysis of the Primary Endpoint – Time to 12-Week CDP	36
3.2.2.5.2	Analysis of Secondary Endpoints	39
3.2.2.6	Reviewer’s Analysis	42
3.2.2.6.1	Analysis of the Primary Endpoint.....	42
3.2.2.6.2	Analysis of Secondary Endpoints	44
3.2.2.6.3	Other Analyses Performed in Response to the Outcomes from Site Inspection	47
3.2.2.6.4	Summary of Reviewer’s Analysis	48
3.3	EVALUATION OF SAFETY	48
4	FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	48

4.1	GENDER, RACE, AGE, AND GEOGRAPHIC REGION	48
4.2	OTHER SPECIAL/SUBGROUP POPULATIONS.....	50
5	SUMMARY AND CONCLUSIONS.....	52
5.1	STATISTICAL ISSUES	52
5.2	COLLECTIVE EVIDENCE	53
5.3	CONCLUSIONS AND RECOMMENDATIONS	54

LIST OF TABLES

Table 1 List of All Studies Included in This Review.....	8
Table 2 Statistical Analysis of the Primary and Secondary Endpoints.....	13
Table 3 Patient Disposition by Study.....	17
Table 4 Summary of Demographic Data by Study.....	18
Table 5 Summary of Patient Disease History and Baseline Disease Characteristics by Study....	19
Table 6 Annualized Protocol-Defined Relapse Rate by Week 96 by Study.....	20
Table 7 Summary of Secondary and Sensitivity Analyses of ARR.....	21
Table 8 Summary of Secondary Efficacy Results	22
Table 9 Sensitivity Analyses of Time to Onset of 12-Week CDP.....	24
Table 10 Assessment of Effect of Discontinuation on Efficacy Results	26
Table 11 Summary of Analysis of 12-Week CDI by Study	26
Table 12 Summary of Results from Analysis of MRI Endpoints	27
Table 13 Statistical Analysis of Primary and Secondary Efficacy Endpoints	32
Table 14 Summary of Reasons for Withdrawal from the Double-blind Treatment Period.....	34
Table 15 Summary of Demographic Data (ITT Population).....	34
Table 16 Summary of baseline disease characteristics (ITT Population).....	35
Table 17 Time to Onset of 12-week CDP during the Treatment Period with Imputation.....	36
Table 18 Sensitivity analyses of the primary endpoint.....	38
Table 19 Time to Onset of 24-week CDP Weeks during the Treatment Period with Imputation	40
Table 20 Summary of Analysis of T25FW.....	41
Table 21 Confirmed Disability Progression without confirmation due to withdrawal.....	43
Table 22 Analysis of 12-week CDP in the Patient Still at Risk at Weeks 24, 48, 72, 96, and 120	43
Table 23 Descriptive statistics and sensitivity analyses of T25FW.....	45
Table 24 Summary of T2 Lesion Volumes at Weeks 24, 48, and 120	46
Table 25 Subgroup Analysis of ARR by Gender and Age, Studies WA21092 and WA21093 ...	49
Table 26 Subgroup Analysis of 12-week CDP by Gender and Age, Studies WA21092 and WA21093.....	49
Table 27 Subgroup Analysis of 12-week CDP by Gender and Age, Study 25046	50
Table 28 Subgroup Analyses of ARR by Region, Baseline EDSS Score and BMI, Studies WA21092 and WA21093	50
Table 29 Subgroup Analyses of 12-week CDP by Region, Baseline EDSS Score and BMI, Studies WA21092 and WA21093.....	51
Table 30 Subgroup Analyses of 12-week CDP by Region, Baseline EDSS Score and BMI, Studies WA25046	52

LIST OF FIGURES

Figure 1 Overview of Study Design - Protocols WA21092 and WA21093.....	9
Figure 2 Hierarchical Order of Key Secondary Efficacy Endpoints	15
Figure 3 Kaplan-Meier Plot for Time to First protocol Defined Relapse for Studies WA21092 and WA21093	21
Figure 4 Kaplan-Meier Plot of Time to Onset of 12-Week CDP for Studies WA21092 and WA21093.....	24
Figure 5 Overview of the Study Design	28
Figure 6 Overview of Patient Disposition	33
Figure 7 Kaplan-Meier Plot of Time to Onset of Confirmed Disability Progression for At Least 12 Weeks during the Double-blind Treatment Period	37

1 EXECUTIVE SUMMARY

Ocrelizumab has been developed by Genentech for the treatment of Relapsing forms of Multiple Sclerosis (RMS) and Primary Progressive Multiple Sclerosis (PPMS). This application is seeking approval for the following indications:

- Ocrelizumab is indicated for the treatment of patients with relapsing forms of multiple sclerosis (b) (4)
- Ocrelizumab is indicated for the treatment of patients with primary progressive multiple sclerosis (PPMS) (b) (4)

Two pivotal trials, WA21092 and WA21093, were conducted to support the claims for treatment of RMS and one pivotal trial, WA25046, was conducted to support the claims for treatment of PPMS.

Trials WA21092 and WA21093 had an identical design: both were randomized, double-blind, double-dummy, parallel-group, active controlled Phase III studies with treatment duration of 96 weeks. The two trials were designed to evaluate the efficacy and safety of ocrelizumab 600 mg every 24 weeks in comparison to interferon beta-1a 44 µg SC in patients with RMS. The primary endpoint was the protocol-defined annualized relapse rate evaluated at 96 weeks. A key secondary endpoint was time to onset of disability progression confirmed at 12 weeks.

Statistically significant treatment difference was shown in annualized relapse rate in both pivotal trials WA21092 and WA21093. Both trials also showed significant treatment difference in 12-week confirmed disability progression.

Trial WA25046 was a randomized, double-blind, parallel-group, multicenter study to evaluate the efficacy and safety of 600 mg ocrelizumab, administered as two 300 mg IV infusions separated by 14 days at a scheduled interval of every 24 weeks, as compared with placebo in adults with PPMS. Patients were treated for a variable duration with the clinical cut-off date (CCOD) occurring after a minimum of 120 weeks and approximately 253 events had been accrued, resulting in at least five treatment doses per patient. The primary endpoint was the time to onset of disability progression that was confirmed at 12 weeks over the treatment period. Secondary endpoints included 24-week confirmed disability progression and change in timed 25-foot walk speed from baseline.

Statistically significant treatment difference was shown in the primary endpoint of time to 12-week confirmed disability progression as well as 24-week confirmed disability progression in trial WA25046.

2 INTRODUCTION

2.1 Overview

The Phase III program for ocrelizumab consisted of two identically designed studies in patients with relapsing forms of multiple sclerosis (RMS) (Studies WA21092 and WA21093) and one study in patients with PPMS (Study WA25046).

This application is seeking approval for the following indications:

- Ocrelizumab is indicated for the treatment of patients with relapsing forms of multiple sclerosis (b) (4).
- Ocrelizumab is indicated for the treatment of patients with primary progressive multiple sclerosis (PPMS) (b) (4).

Currently, no treatment has been demonstrated to significantly slow the progression of disability in patients with PPMS.

The two trials (WA21092 and WA21093) for treatment of patients with RMS had an identical design. Both were randomized, double-blind, double-dummy, parallel-group, active controlled study with treatment duration of 96 weeks. The two trials were to assess whether the efficacy of ocrelizumab 600 mg every 24 weeks (given as two infusions of 300 mg on Days 1 and 15 of the first 24-week treatment period, and as a single infusion of 600 mg subsequently) was superior to interferon beta-1a 44 µg SC as measured by the protocol-defined annualized relapse rate at 96 weeks.

Trial WA25046 was designed to evaluate the efficacy and safety of 600 mg ocrelizumab, administered as two 300 mg IV infusions separated by 14 days at a scheduled interval of every 24 weeks, as compared with placebo in adults with PPMS. The trial was a randomized, double-blind, placebo-controlled, parallel-group, multicenter study. Patients were treated for a variable duration with the clinical cut-off date (CCOD) occurring after a minimum of 120 weeks and approximately 253 events had been accrued, resulting in at least five treatment doses per patient. The primary endpoint was the time to onset of disability progression that was confirmed at 12-week over the treatment period.

The following table presents a summary of the studies included in this review.

Table 1 List of All Studies Included in This Review

	Phase and Design	Treatment Period	Comparator	# of Subjects per Arm	Study Population
WA21092	Phase 3, randomized, double-blind, 2-arm, active-controlled	96 weeks	IFNB-1a	Ocrelizumab: 411 IFNB-1a: 410	Patients with RMS
WA21093	Phase 3, randomized, double-blind, 2-arm, active-controlled	96 weeks	IFNB-1a	Ocrelizumab: 418 IFNB-1a: 417	Patients with RMS
WA25046	Phase 3, randomized, double-blind, 2-arm, active-controlled	Various length with minimum treatment of 120 weeks	Placebo	Ocrelizumab: 488 Placebo: 244	Patients with PPMS

Source: Reviewer's summary

2.2 Data Sources

All documents reviewed for this BLA supplement submission are in electronic form. The path to the original submission on 4/28/2016 is <\\CDSESUB1\evsprod\BLA761053\0001>.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

No major issues were identified in the submission of data and study documents.

3.2 Evaluation of Efficacy

3.2.1 Evaluation of Efficacy for Studies WA21092 and WA21093

3.2.1.1 Study Design

Studies WA21092 and WA21093 had an identical design. The primary objective of the two studies was to assess whether the efficacy of ocrelizumab 600 mg every 24 weeks was superior to interferon beta-1a 44 µg SC as measured by the protocol-defined annualized relapse rate by 96 weeks in patients with RMS.

The two studies WA21092 and WA21093 were randomized, double-blind, double-dummy, parallel-group, active controlled Phase III study with treatment duration of 96 weeks.

Eligible patients were randomized 1:1 to a regimen of either ocrelizumab 600 mg by IV infusion every 24 weeks (given as two infusions of 300 mg on Days 1 and 15 of the first 24-week treatment period, and as a single infusion of 600 mg subsequently) or interferon beta-1a 44 µg by SC injection 3 times per week for 96 weeks. Randomization was stratified by region (United States versus rest of the world [ROW]) and baseline EDSS (<4.0 versus ≥ 4.0).

During the 96-week treatment patients were required to attend 10 scheduled assessment visits. Additionally, unscheduled visits were made for the assessment of potential relapses. A schematic of the study design is presented in Figure 1.

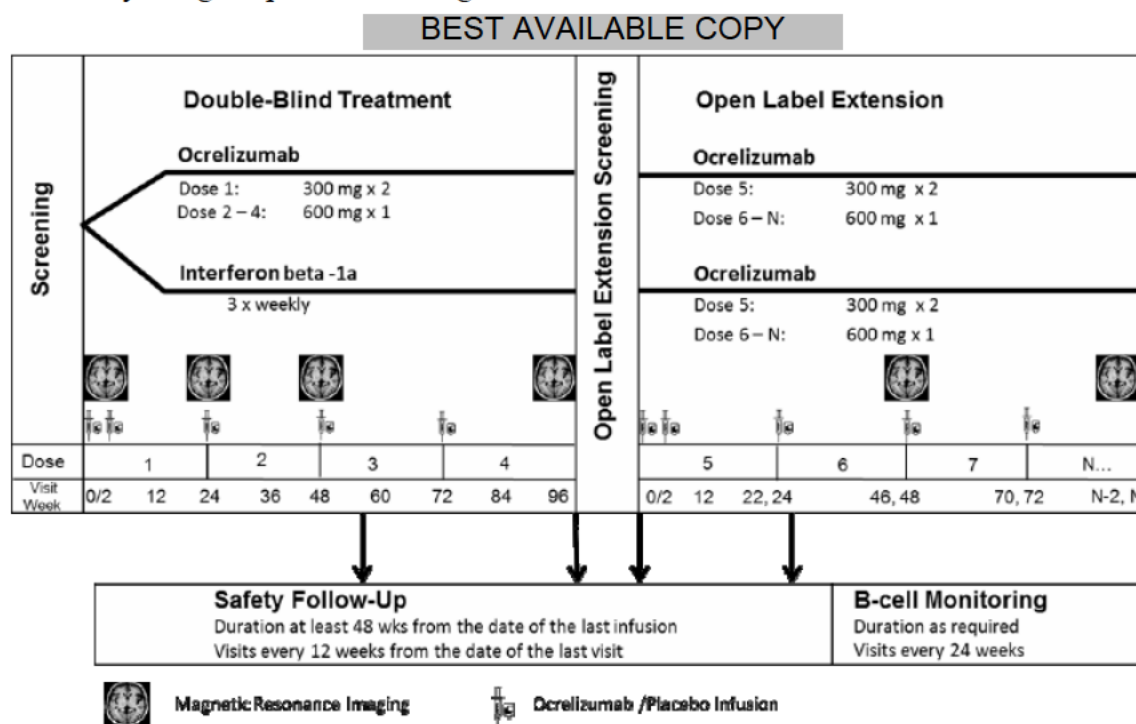


Figure 1 Overview of Study Design - Protocols WA21092 and WA21093

Source: CSR

Patients who prematurely discontinued study medication during the 96-week double-blind treatment period attended a withdrawal visit for the purpose of collecting end-of-treatment assessments. These patients subsequently entered the safety follow-up (SFU). Patients who completed the 96-week double-blind (DB) treatment had an option to enter a single group, active treatment open-label extension (OLE) providing they fulfilled the eligibility criteria.

The two studies enrolled patients of ages 18-55 years with diagnosis of MS in accordance with the revised McDonald criteria (2010) and had at least 2 documented clinical attacks within the last 2 years prior to screening, or one clinical attack in the year prior to screening, with neurological stability for ≥ 30 days prior to both screening. Patients had baseline EDSS from 0 to

5.5 (inclusive) at screening and documented MRI of brain with abnormalities consistent with MS prior to screening.

Patients who had diagnosis of primary progressive MS were excluded from the two studies.

3.2.1.2 Study Endpoints

The two studies WA21092 and WA21093 had identical primary and secondary efficacy endpoints.

The primary efficacy endpoint was the annualized rate of protocol-defined relapse at 96 weeks.

The secondary efficacy endpoints were listed in descending order of importance, as described below.

- The time to onset of confirmed disability progression (CDP) for at least 12 weeks (12-week CDP), with the initial event of neurological worsening occurring during the 96-week treatment period
- The total number of T1 Gd-enhancing lesions as detected by brain MRI at Weeks 24, 48, and 96
- The total number of new and/or enlarging T2 hyperintense lesions as detected by brain MRI at Weeks 24, 48, and 96
- The proportion of patients who have confirmed disability improvement (CDI) for at least 12 weeks (12-week CDI), with the initial event of neurological improvement occurring during the 96-week treatment period
- The time to onset of CDP for at least 24 weeks (24-week CDP), with the initial event of neurological worsening occurring during the 96-week treatment period
- The total number of new T1 hypointense lesions at Weeks 24, 48, and 96
- The change in MSFC score from baseline to Week 96
- The percentage change in brain volume as detected by brain MRI from Week 24 to Week 96
- The change in SF-36 PCS Score from baseline to Week 96
- The proportion of patients who have No evidence of Disease Activity (NEDA) by Week 96

3.2.1.2.1 Definition of Efficacy Endpoints

Protocol Defined Relapse

The primary efficacy endpoint was the protocol-defined annualized relapse rate (ARR) at 96 weeks. A protocol-defined relapse (PDR) was described by all of the following criteria:

- occurrence of new or worsening neurological symptoms attributable to MS
- symptoms persisting for >24 hours

- symptoms not attributable to confounding clinical factors (e.g., fever, infection, injury, adverse reactions to medications)
- symptoms immediately preceded by a stable or improving neurological state for at least 30 days
- symptoms accompanied by objective neurological worsening consistent with an increase of at least half a step on the EDSS scale, or 2 points on one of the appropriate FSS, or 1 point on two or more of the appropriate FSS; the change had to affect the selected FSS (i.e., pyramidal, ambulation, cerebellar, brainstem, sensory, or visual).

Confirmed Disability Progression Definition

Disability progression was defined as an increase from the baseline in the EDSS score of ≥ 1.0 point if the baseline score was ≤ 5.5 or at least 0.5 point if the baseline score was > 5.5 .

Disability progression was considered confirmed when the increase in the EDSS was confirmed at a regularly scheduled visit at least 12 weeks (for 12-week CDP) or 24 weeks (for 24-week CDP) after the initial documentation of neurological worsening.

Initial progression could occur at any visit during the 96-week treatment period. Confirmation of disability progression had to occur at a regularly scheduled visit during the treatment phase or in the SFU phase or OLE phase after study-drug discontinuation that was at least 12 weeks (or 24 weeks) after the initial progression. If a patient had a missing EDSS at the scheduled visit occurring at least 84 days after an initial progression or the scheduled visit occurred several days before the 84-day window after an initial progression (e.g., the visit window is ± 4 days), confirmation of the disability progression had to be on the basis of the assessment at the next scheduled visit.

Disability progression could not be confirmed unless the EDSS values met the minimum change required for progression. The non-confirmatory EDSS assessments (if any) between the initial disability progression and the confirmation of disability progression had to be at least as high as the minimum change required for progression.

Patients who, according to the above definition, did not have onset of disability progression by the Week-96 visit, by the time of the early discontinuation of treatment, or were lost to follow up were to be censored at the date of the last EDSS assessment during the 96-week treatment period.

Confirmed Disability Improvement Definition

Disability improvement was assessed only for the subgroup of patients with a baseline EDSS score of ≥ 2.0 . It was defined as a reduction from baseline in EDSS score of ≥ 1.0 point if the baseline score was ≥ 2 and ≤ 5.5 and ≥ 0.5 point if the baseline EDSS score was > 5.5 .

All patients without confirmed disability improvement were counted as not improved, independent of follow-up time.

No Evidence of Disease Activity (NEDA) Definition

NEDA was defined only for patients with a baseline EDSS score ≥ 2.0 . Patients who completed the 96-week treatment period were considered as having evidence of disease activity if at least one PDR, a CDP event or at least one MRI scan showing MRI activity (defined as Gd-enhancing T1 lesions, or new or enlarging T2 lesions) was reported during the 96-week treatment period; otherwise the patient was considered as having NEDA. Patients who discontinued treatment early with at least one event before early discontinuation were considered as having evidence of disease activity.

3.2.1.3 Statistical Methodologies

3.2.1.3.1 General Consideration

All randomized patients were to be included in the intent-to-treat (ITT) population according to their randomized treatment. All efficacy analyses were to be performed using the ITT population.

The baseline EDSS was defined as the average EDSS score from the screening and baseline visit. For the analysis of mean change of EDSS, disability progression, and disability improvement, the exact value of the average EDSS score was to be used. However, for the stratification factor and for the randomization algorithm, the average was rounded up to the next EDSS score in order to assign a valid score (e.g., round 3.75 to 4.0 and round 5.25 to 5.5). If one of the EDSS scores from the screening or baseline visit was missing, the other was to be used. If EDSS score was missing at both screening and baseline visit, the patient was not to be randomized.

All primary and secondary efficacy endpoints were to be analyzed, adjusted or stratified by geographical region (United States vs. ROW) and baseline EDSS (< 4.0 vs. ≥ 4.0).

The null hypothesis to be tested was that there was no statistically significant difference between ocrelizumab group (OCR group) and the interferon beta-1a group (IFN group). All statistical hypotheses for the primary and secondary endpoints and treatment comparisons were to be tested at the 5% significance level ($\alpha = 0.05$) against two-sided alternatives.

The following table is a summary of statistical analyses of the primary and secondary endpoints. Details of the analyses are described in the sub-sections.

Table 2 Statistical Analysis of the Primary and Secondary Endpoints

Endpoint	Statistical Model	Stratification/adjusting factors
Primary Endpoint		
ARR	NBR (offset variable = log-transformed exposure time)	Baseline EDSS (<4.0 vs. ≥4.0), geographical region (US vs. ROW)
Secondary Endpoints		
Disability		
Time to onset of CDP for at least 12 weeks / 24 weeks	Log-rank test, Cox regression (for estimation of hazard ratio (HR))	Baseline EDSS (<4.0 vs. ≥4.0), geographical region (US vs. ROW)
Proportion of patients who have CDI for at least 12 weeks	CMH χ^2	Baseline EDSS (<4.0 vs. ≥4.0), geographical region (US vs. ROW)
Mean change from baseline in MSFC at Week 96	MMRM	Baseline MSFC, baseline EDSS (<4.0 vs. ≥4.0), geographical region (US vs. ROW)
Brain MRI		
T1Gd-enhancing lesions per MRI scan	NBR (offset variable = log-transformed number of brain MRI scans received)	Baseline T1-Gd-enhanced lesion (present vs. absent), baseline EDSS score (<4.0 vs ≥ 4.0), geographical region (US vs. ROW)
New and/or enlarging T2 hyperintense lesions per MRI scan	NBR (offset variable = log-transformed number of brain MRI scans received)	Baseline T2 hyperintense lesion count, baseline EDSS score (<4.0 vs ≥ 4.0), geographical region (US vs. ROW)
New T1 hypointense lesions per MRI scan	NBR (offset variable = log-transformed number of brain MRI scans received)	Baseline T1-hypointense lesions count, baseline EDSS score (<4.0 vs ≥ 4.0), geographical region (US vs. ROW)
Mean % change from Week 24 in brain volume at Week 96	MMRM	Week 24 brain volume, baseline EDSS (<4.0 vs. ≥4.0), geographical region (US vs. ROW)
Disease Activity		
Proportion of patients with NEDA at Week 96	CMH χ^2	Baseline EDSS (<4.0 vs. ≥4.0), geographical region (US vs. ROW)
Quality of Life		
Mean change from baseline in SF-36 PCS at Week 96	MMRM	Baseline SF-36 PCS, baseline EDSS (<4.0 vs. ≥4.0), geographical region (US vs. ROW)

Source: CSR

3.2.1.3.2 Analyses of the Primary Endpoints

The ARR by 96 weeks was to be analyzed using a negative binomial model, adjusting for region (United States versus ROW) and baseline EDSS (< 4.0 versus ≥ 4.0). In order to account for different study treatment durations among patients, the exposure of each patient was to be taken into account by including log-transformed exposure time in the model as an “offset” variable for appropriate computation of relapse rate.

Sensitivity analyses of the primary endpoint included:

- ARR calculated including all protocol-defined relapses occurring during the treatment period or the SFU, up to 96 weeks after randomization. This analysis used all available data and estimated the treatment effect up to 96 weeks irrespective of whether patients were on study treatment or not.
- Per-protocol population
- Safety population
- Adjustment by additional covariates (number of relapses within 2 years prior to study entry, baseline Gd lesions [presence vs. absence], prior MS treatment, age [<40 , ≥ 40 years])
- Poisson model with the same covariates as in the negative binomial regression
- Use of two different methods for handling missing data (patients prematurely withdrawn from treatment), to explore the potential influence of informative dropouts on the results of the primary efficacy analysis (only data from the treatment period was considered):
 - Multiple imputation: for patients that discontinued early during the double-blind, double-dummy treatment period without any PDR in the 30 days prior to discontinuation, 50% of patients were randomly assigned an event of relapse on day of discontinuation, and 50% were censored on day of discontinuation;
 - With imputation: patients that discontinued early during the double-blind, double-dummy treatment period without any PDR in the 30 days prior to discontinuation, were counted as having had a relapse on day of discontinuation.

3.2.1.3.3 Analysis of Secondary Endpoints

Following the analysis of primary endpoint, secondary endpoints were to be analyzed in the hierarchical order as shown in Figure 2, all at $\alpha = 0.05$ level.

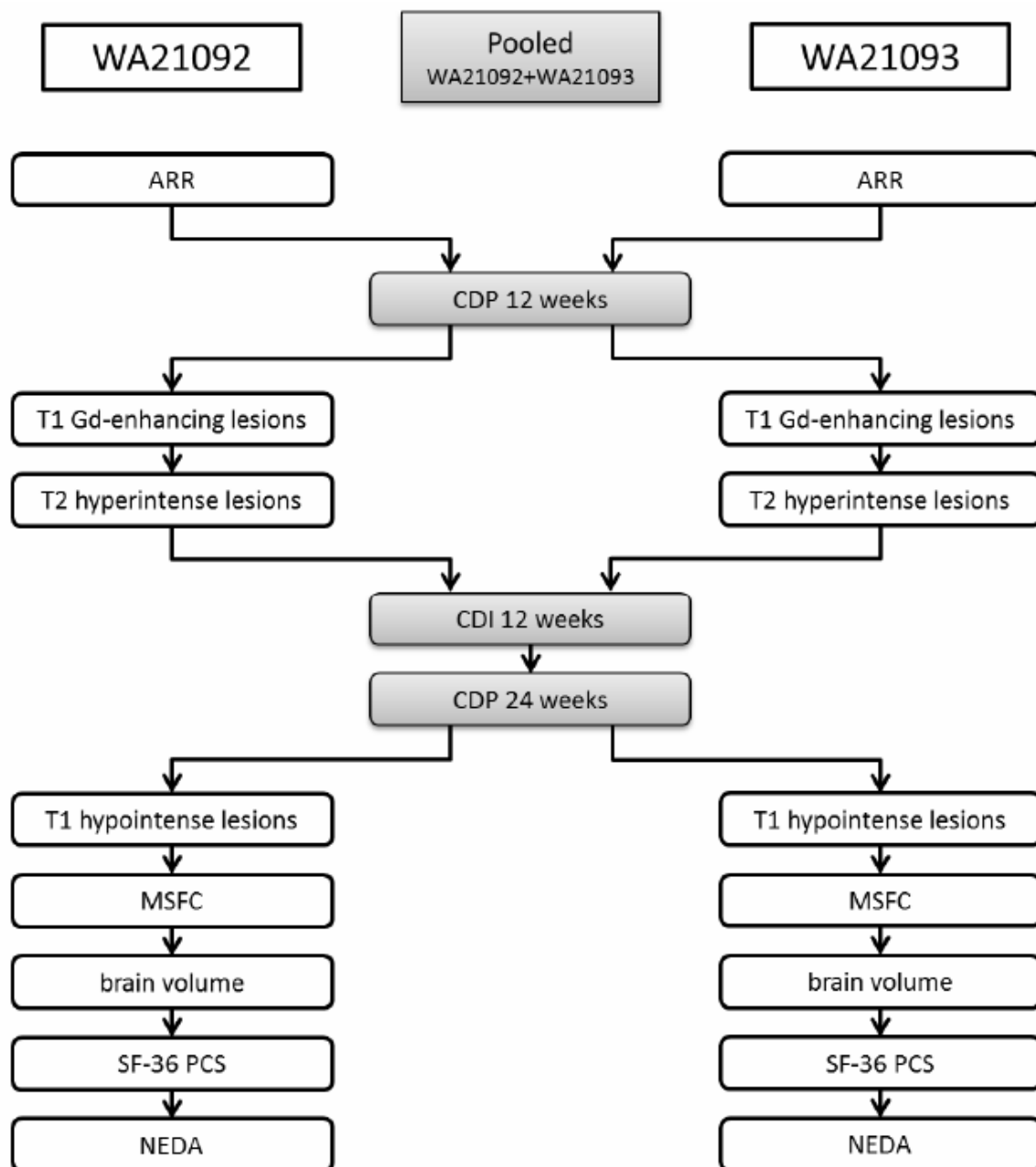


Figure 2 Hierarchical Order of Key Secondary Efficacy Endpoints

Source: CSR

For disability related endpoints (CDP at 12 weeks, CDI at 12 weeks and CDP at 24 weeks) data from the 2 studies WA01292 and WA01293 were to be pooled together in the analysis as the two studies were not sufficiently powered to detect treatment difference by each individual study. Only data collected during the 96-week treatment period were to be used for pooled analysis of disability, although confirmation of disability progression during the 96-week treatment period could occur during a subsequent study phase (i.e., during the SFU phase or the OLE phase).

Analysis of Disability Progression

Time to confirmed disability progression in the ocrelizumab group and in the interferon beta-1a 44 µg SC group was to be compared using a two-sided log-rank test stratifying by geographical region (United States vs. ROW), and baseline EDSS (< 4.0 vs. ≥ 4.0). The proportion of patients with confirmed disability progression was to be estimated using Kaplan-Meier methodology. The overall hazard ratio was to be estimated using a stratified Cox regression model with the same stratification factors used in the log-rank test above.

Analysis of T1 Gd-Enhancing Lesions

The total number of T1 Gd-enhanced lesions was to be calculated as the sum of the individual number of T1 Gd-enhanced lesions at Weeks 24, 48, and 96. The same approach used for the primary analysis of ARR was to be applied with the following exception:

- In order to account for patients receiving varying numbers of brain MRI scans during the study, the log-transformed number of brain MRI scans received was to be included in the model as an “offset” variable for appropriate computation.
- Baseline Gd lesion (present or not) was to be added as an additional covariate in the negative binomial model.

Analysis of New or Enlarging T2 Lesions

The total number of new and/or enlarging T2 lesions was to be calculated as the sum of the individual number of new and/or enlarging lesions at Weeks 24, 48, and 96. The same approach used for the total number of T1 Gd-enhancing lesions was to be applied with the exception that baseline T2 lesion count was to be added as an additional covariate in the negative binomial model.

Analysis of T1 Hypo-Intense Lesions

The total number of new T1-hypo-intense lesions (chronic black holes) was to be calculated as the sum of the individual number of new lesions at Weeks 24, 48, and 96. The same approach used for the total number of T1 Gd-enhancing lesions was to be applied with the exception that baseline T1 hypo-intense lesion count was to be added as an additional covariate in the negative binomial model.

Analysis of Confirmed Disability Improvement

This endpoint was to be analyzed only for the subgroup of patients with a baseline EDSS score ≥ 2.0.

The proportions in treatment groups was to be compared using the Cochran-Mantel-Haenszel (CMH) χ^2 test stratified by geographical region (United States vs. ROW), and baseline EDSS score (< 4.0 vs. ≥ 4.0). The two studies WA21092 and WA21093 were to be combined.

Other Secondary Efficacy Endpoints

The endpoints of MSFC, sub-scores of MSFC and brain volume are not considered validated and/or clinically meaningful endpoints, and are therefore not presented in this review.

3.2.1.4 Study Results of Protocols WA21092 and WA21093

3.2.1.4.1 Patient Disposition

A summary of patient disposition for Studies WA21092 and WA21093 is presented in Table 3.

Table 3 Patient Disposition by Study

	WA21092		WA21093	
	IFN beta-1a	OCR 600 mg	IFN beta-1a	OCR 600 mg
Randomized	411	410	418	417
Received Treatment	409	408	417	0
Discontinued	71 (17.3%)	44 (10.7%)	98	57
Prior to treatment	2	2	1	0
After dosing	69	42	97	57
Reason for Discontinue				
AE	25	13	25	16
Death	1	0	1	1
Lack of Efficacy	12	8	15	6
Lost to Follow-Up	1	1	10	6
Non-Compliance	5	0	2	4
Physician Decision	0	1	0	1
Pregnancy	2	3	3	0
Protocol Violation	1	2	1	1
Subject Decision	13	8	25	12
Other	11	8	16	10

Source: CSR

In Study WA21092, a total of 1051 patients were screened and 821 patients were enrolled at 141 investigational sites across 32 countries. Of these 821 patients, 411 patients were randomized to IFN group and 410 patients were randomized to OCR group.

Two patients in each treatment group withdrew from the study before receiving any treatment. A total of 115 patients (14%) withdrew from the double-blind treatment period in Study WA21092 prior to completing Week 96. A greater proportion of patients in the IFN group (17%) withdrew prematurely from treatment during the double-blind treatment period compared with the OCR group (11%) including 2 patients in each group who withdrew before receiving study drug.

In Study WA21093, a total of 1045 patients were screened and 835 patients were enrolled at 166 investigational sites across 24 countries. Of the enrolled patients, 418 patients were randomized to IFN group and 417 patients were randomized to OCR group. One patient in the IFN group withdrew prior to receiving any study treatment due to pregnancy.

In total, 155 patients (19%) withdrew from the double-blind treatment period prior to completing Week 96. A greater proportion of patients in the IFN group (23%) withdrew prematurely from the double-blind treatment period compared with the OCR group (14%); including 1 patient in the IFN group who withdrew before receiving study drug.

3.2.1.4.2 Patient Demographics

Demographic characteristics of the ITT population were well balanced across the treatment groups in both WA21092 and WA21093. The majority of patients was female (about 66%) and patients were predominantly white (about 90%) with median ages of 37 to 38 years in the treatment groups. Patients were mainly from Europe (EU, Switzerland and Norway) or North America (USA and Canada) and Australia.

Table 4 Summary of Demographic Data by Study

	WA21092		WA21093	
	IFN beta-1a N=411	OCR 600 mg N=410	IFN beta-1a N=418	OCR 600 mg N=417
Age				
Mean (Median)	36.9 (37.0)	37.1 (38.0)	37.4 (38.0)	37.2 (37.0)
Min – Max	18 – 55	18 – 56	18 – 55	18 – 55
N (%) < 40	243 (59.1%)	244 (59.5%)	241 (57.7%)	252 (60.4%)
N (%) ≥ 40	168 (40.9%)	166 (40.5%)	177 (42.3%)	165 (39.6%)
Sex, N (%)				
Male	139 (33.8%)	140 (34.1%)	138 (33.0%)	146 (35.0%)
Female	272 (66.2%)	270 (65.9%)	280 (67.0%)	271 (65.0%)
Race, N (%)				
White	375 (91.2%)	375 (91.5%)	382 (91.4%)	368 (88.2%)
African American	12 (2.9%)	19 (4.6%)	20 (4.8%)	21 (5.0%)
Other	24 (5.8%)	16 (3.9%)	16 (3.8%)	28 (6.7%)
Weight (kg)				
Mean (Median)	75.9 (74.0)	74.6 (72.0)	75.0 (72.5)	75.9 (73.8)
Min – Max	43.0 – 137.4	41.0 – 170.0	42.1 – 163.6	38.0 – 135.0
Region, N (%)				
USA	105 (25.5%)	105 (25.6%)	114 (27.3%)	112 (26.9%)
ROW	306 (74.5%)	305 (74.4%)	304 (72.7%)	305 (73.1%)
EU/Switzerland/Norway	204 (49.6%)	211 (51.5%)	182 (43.5%)	187 (44.8%)
Latin America	35 (8.5%)	26 (6.3%)	26 (6.2%)	19 (4.6%)
Non-EU/Israel/Africa	64 (15.6%)	68 (16.6%)	53 (12.7%)	58 (13.9%)
USA/Canada/Australia	108 (26.3%)	105 (25.6%)	157 (37.6%)	153 (36.7%)

Source: Reviewer's Summary

3.2.1.4.3 Patient Baseline Disease Characteristics

In both studies, the two treatment groups had similar MS disease histories and baseline disease characteristics (Table 5).

Table 5 Summary of Patient Disease History and Baseline Disease Characteristics by Study

	WA21092		WA21093	
	IFN beta-1a N=411	OCR 600 mg N=410	IFN beta-1a N=418	OCR 600 mg N=417
Years since MS symptom onset Mean (Median)	6.3 (4.6)	6.7 (4.9)	6.7 (5.1)	6.7 (5.2)
Number of relapses prior year Mean (Median)	1.3 (1.0)	1.3 (1.0)	1.3 (1.0)	1.3 (1.0)
Number of relapses past 2 years Mean (Median)	1.7 (2.0)	1.8 (2.0)	1.8 (2.0)	1.8 (2.0)
EDSS Mean (median) N (%) < 4 N (%) ≥ 4	2.8 (2.5) 318 (77.6%) 92 (22.4%)	2.9 (2.5) 314 (76.6%) 96 (23.4%)	2.8 (2.5) 309 (73.9%) 109 (26.1%)	2.8 (2.5) 315 (71.5%) 102 (24.5%)
Number of Gd+ T1 lesions Mean (Median)	1.9 (0.0)	1.7 (0.0)	2.0 (0.0)	1.8 (0.0)
Number of T1 Hypointense lesions Mean (Median)	32.9 (21.0)	33.1 (21.0)	32.8 (21.0)	31.6 (20.0)
Volume of T2 lesions (cm ³) Mean (Median)	9.7 (6.2)	10.8 (5.7)	10.6 (6.2)	10.7 (5.3)
Number of T2 lesions Mean (Median)	51.1 (41.0)	51.0 (40.5)	51.0 (45.0)	49.3 (39.0)
Brain volume (cm ³) Mean (Median)	1499.2 (1503.6)	1500.9 (1498.8)	1501.1 (1506.5)	1503.9 (1510.5)

Source: Reviewer's Summary

3.2.1.5 Efficacy Results of WA21092 and WA21093 Reported by the Sponsor

The sponsor reported that in both studies WA21092 and WA21093, all patients randomized to treatment (821 in WA21092 and 835 in WA21093) were included in the ITT population.

3.2.1.5.1 Annualized Relapse Rate (ARR) – Primary Endpoint

Both studies met the primary objectives with respect to efficacy: at 96 weeks, treatment with ocrelizumab had a statistically significant reduction on the ARR by 46.4% for WA21092 and 46.8% in WA21093 compared with interferon beta-1a ($p < 0.0001$ for both studies).

The results from the primary analysis for the two studies were similar. In Study WA21092, the adjusted ARR at 96 weeks was 0.292 for the IFN group and 0.156 for the OCR group. In Study WA21093, the adjusted ARR at 96 weeks was 0.290 for the IFN group and 0.155 for the OCR group.

Table 6 Annualized Protocol-Defined Relapse Rate by Week 96 by Study

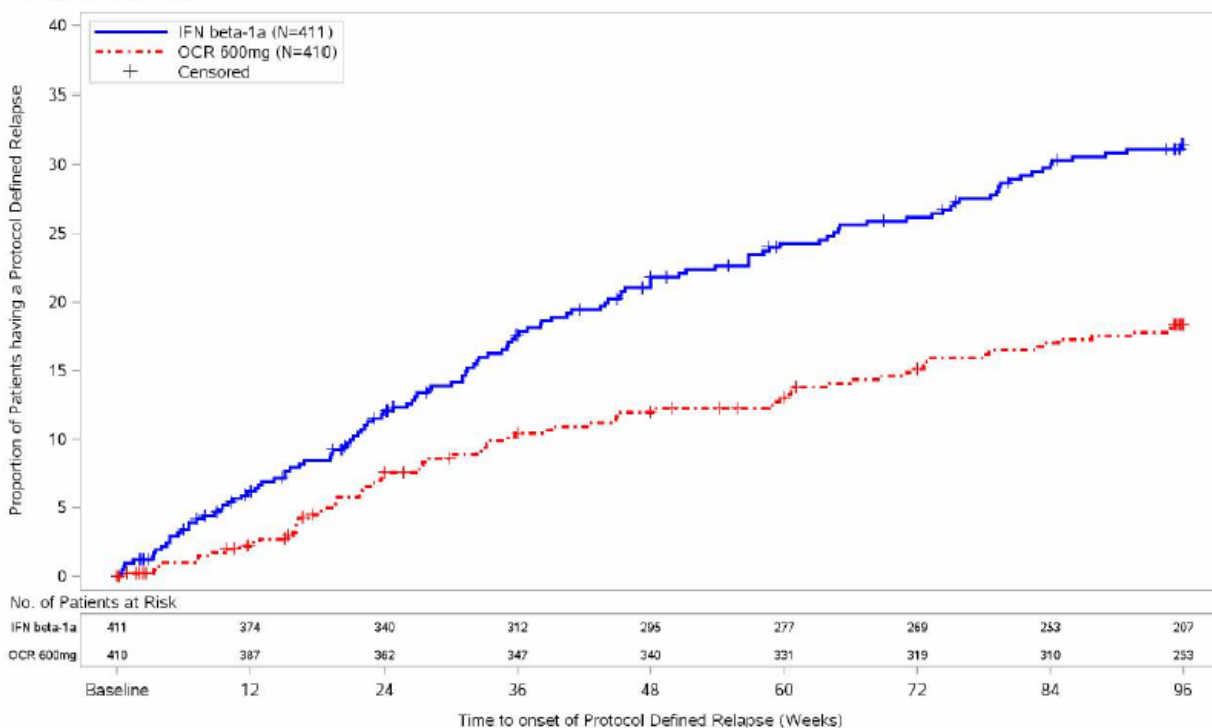
	WA21092		WA21093	
	IFN beta-1a (N=411)	OCR 600 mg (N=410)	IFN beta-1a (N=418)	OCR 600 mg (N=417)
Total number of relapses	166	96	168	98
Total patient-years followed	678.1	706.3	661.0	709.5
Unadjusted ARR*	0.245	0.136	0.254	0.138
Adjusted ARR**	0.292	0.156	0.290	0.155
95% CI of adjusted ARR	(.235, 0.361)	(0.122, 0.200)	(0.234, 0.361)	(0.121, 0.198)
ARR rate ratio		0.536		0.532
95% CI of rate ratio		(0.400, 0.719)		(0.397, 0.714)
p-value		<.0001		<.0001

*The total number of relapses for all patients in the treatment group divided by the total patient-years of exposure to that treatment.

** Adjusted by baseline EDSS (<4.0 vs. ≥4.0) and geographic region (US vs. ROW).

Source: CSR

The Kaplan-Meier curves for time to first protocol-defined relapse for the two studies are presented in the following figure. The log-rank test for the time to first PDR yielded a p-value of < 0.0001 for both studies. The hazard ratio from the Cox model was 0.536 for both Studies WA21092 and WA21093.

Protocol: WA21092

Protocol: WA21093

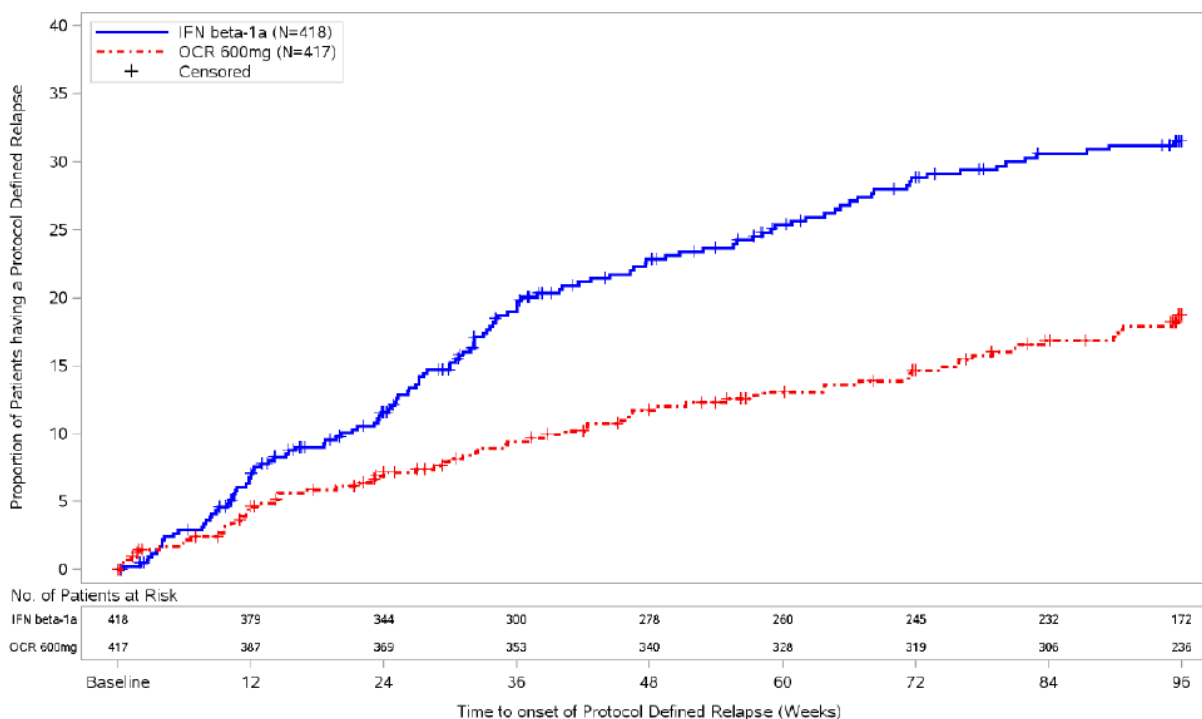


Figure 3 Kaplan-Meier Plot for Time to First protocol Defined Relapse for Studies WA21092 and WA21093

Secondary and sensitivity analyses were performed by the sponsor and results appeared to be consistent with the one from the primary analysis. The following table presents the results reported by the sponsor.

Table 7 Summary of Secondary and Sensitivity Analyses of ARR

	WA21092		WA21093	
	Adjusted Rate Ratio (95% CI)	p-value	Adjusted ARR Ratio(95% CI)	p-value
Including relapses in SFU	0.533 (0.416, 0.734)	<.0001	0.518 (0.389, 0.691)	<.0001
PP Population	0.514 (0.380, 0.696)	<.0001	0.528 (0.391, 0.712)	<.0001
With additional covariates ^a	0.541 (0.405, 0.723)	<.0001	0.547 (0.409, 0.732)	<.0001
Poisson model	0.522 (0.432, 0.706)	<.0001	0.545 (0.427, 0.697)	<.0001
50% imputation for missing ^b	0.538 (0.402, 0.718)	N/A	0.516 (0.390, 0.684)	N/A
100% imputation for missing ^c	0.537 (0.411, 0.701)	<.0001	0.507 (0.397, 0.647)	<.0001

a. Number of relapses occurring within 2 years prior to study entry, presence or absence of baseline Gd lesions, prior MS treatment, age (< 40, ≥ 40).

b. For patients who discontinued early during the double-blind, double-dummy treatment period without any protocol-defined relapse in the 30 days prior to discontinuation, 50% of patients were assigned an event of relapse on day of discontinuation, and 50% were censored on day of discontinuation

c. As in footnote b, except 100% patients who discontinued counted as having had a relapse on day of discontinuation.

Source: CSR

3.2.1.5.2 Secondary Endpoints

A summary of results from analysis of secondary endpoints is presented in Table 8 followed by details of analysis of each of the secondary endpoints.

A hierarchical approach was taken in testing to control the type-I error under 0.05. The sequential testing procedure stopped at MSFC in Study WA21092 and it stopped at brain volume in Study WA21093 when the significance was lost.

Table 8 Summary of Secondary Efficacy Results

	WA21092		WA21093	
	IFN beta-1a (N=411)	OCR 600 mg (N=410)	IFN beta-1a (N=418)	OCR 600 mg (N=417)
12-week CDP N (%) of patients with events Hazard ratio p-value (log-rank)	N=411 50 (12.17)	N=410 31 (7.56) 0.57 0.0139	N=418 63 (15.07)	N=417 44 (10.55) 0.63 0.0169
12-week CDP Pooled (primary) % of patients with events Hazard ratio p-value (log-rank)	N=829 15.18	N=827 9.75 0.60 0.0006	N=829 15.18	N=827 9.75 0.60 0.0006
T1 Gd-enhancing lesions Mean per scan Rate ratio p-value	N=377 0.286	N=388 0.016 0.058 <.0001	N=375 0.416	N=389 0.021 0.051 <0.0001
New/enlarging T2 lesions Mean per scan Rate ratio p-value	N=378 0.982	N=390 0.323 .229 <.0001	N=376 1.904	N=390 0.325 0.171 <0.0001
12-week CDI Proportion of improved Relative risk p-value	N=306 12.42	N=310 20.00 1.61 0.0106	N=308 18.83	N=318 21.38 1.14 0.4019
12-week CDI pooled Proportion of improved Relative risk p-value	N=614 15.64	N=628 20.70 1.33 0.0194	N=614 15.64	N=628 20.70 1.33 0.0194
24-week CDP N (%) of patients with events Hazard ratio p-value (log-rank)	N=411 39 (9.49)	N=410 24 (5.85) 0.57 0.0278	N=418 13.63	N=417 8.60 0.63 0.0370
24-week CDP pooled % of patients with events Hazard ratio p-value (log-rank)	N=829 12.03	N=827 7.58 0.60 0.0025	N=829 12.03	N=827 7.58 0.60 0.0025
New T1 hypointense lesions Mean per scan Rate ratio p-value	N=377 0.982	N=388 0.420 0.428 <0.0001	N=375 1.255	N=389 0.449 0.357 <0.0001

MSFC	N=308	N=322	N=269	N=308
Mean change from baseline	0.174	0.213	0.169	0.276
Mean difference		0.039		0.107
p-value		0.3261		0.0040
Brain Volume	N=267	N=281	N=259	N=287
Mean % change from week 24	-0.741	-0.572	-0.750	-0.638
Mean difference		0.168		0.112
p-value		0.0042		0.0900

Source: Reviewer's summary from CSR

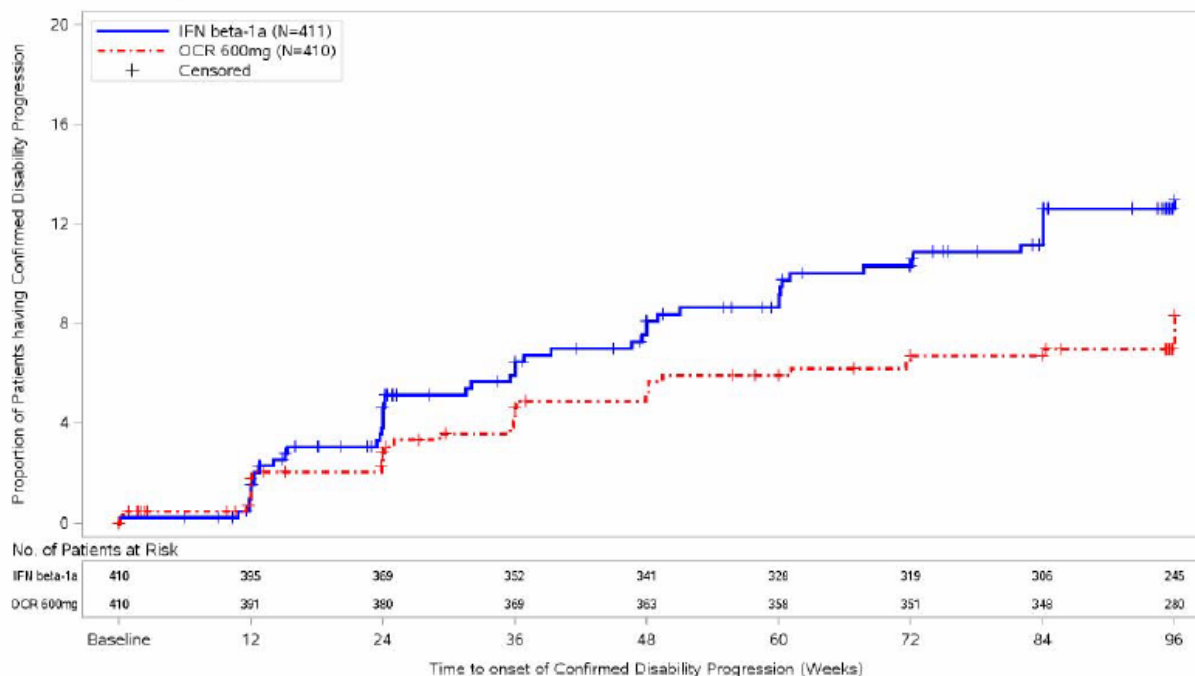
12-Week and 24-week Confirmed Disability Progression

In the **pooled analysis** from Studies WA21092 and WA21093, treatment with ocrelizumab showed a 40% risk reduction in both 12-week CDP (hazard ratio=0.60; p=0.0006) and 24-week CDP (hazard ratio=0.60; p=0.0025) compared with interferon beta-1a.

For each **individual study**, a risk reduction of 43% in both 12-week CDP and 24-week CDP on ocrelizumab was seen in Study WA21092. In Study WA21093, the risk reductions for 12-week CDP and 24-week CDP were both 37%.

The Kaplan-Meier plots of time to onset of 12-week CDP for Studies WA21092 and WA21093 are presented in the following figure.

Protocol: WA21092



Protocol: WA21093

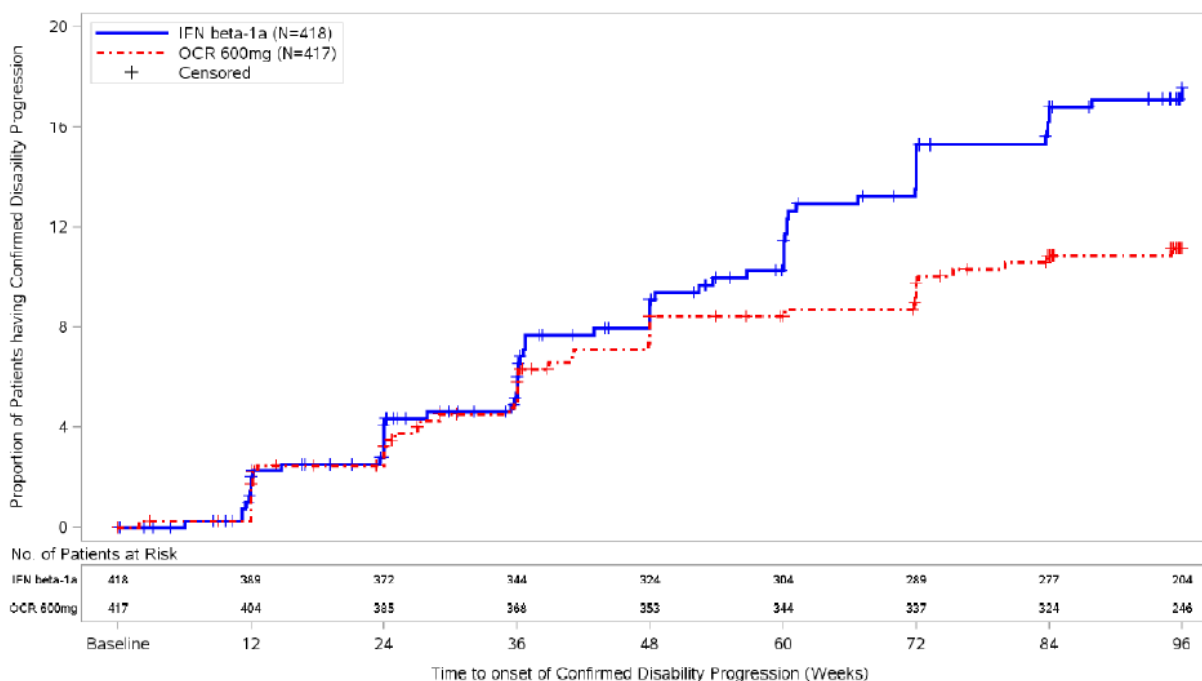


Figure 4 Kaplan-Meier Plot of Time to Onset of 12-Week CDP for Studies WA21092 and WA21093
Source: CSR

Similar sensitivity analyses as to the primary endpoint were performed and the results are presented in the following table.

Table 9 Sensitivity Analyses of Time to Onset of 12-Week CDP

	WA21092		WA21093	
	HR (95% CI)	p-value	HR (95% CI)	p-value
Including initial event in SFU up to 96 weeks ^a	0.60 (0.39, 0.94)	0.0227	0.60 (0.41, 0.88)	0.0088
PP Population	0.56 (0.35, 0.88)	0.0112	0.62 (0.42, 0.93)	0.0183
Adjusted for additional covariates ^b	0.56 (0.35, 0.87)	0.0139	0.65 (0.44, 0.96)	0.0169
Stratified for additional covariates ^c	0.47 (0.28, 0.78)	0.0030	0.56 (0.36, 0.88)	0.0104
50% imputation for missing ^d	0.58 (0.37, 0.90)	N/A	0.63 (0.43, 0.93)	N/A
100% imputation for missing ^e	0.57 (0.37, 0.89)	0.0125	0.64 (0.44, 0.92)	0.0145

- Post-hoc analysis defined after database lock.
- Additional baseline factors: the number of relapses occurring within a 2-year period prior to study entry, baseline presence of T1 Gd-enhancing lesions (present or absent), prior MS treatment and age (<40, ≥ 40 years).
- Stratified by Study, Geographical Region (US vs. ROW), Baseline EDSS (<4.0 vs. ≥ 4.0), the number of relapses occurring within a 2-year period prior to study entry, baseline presence of T1 Gd-enhancing lesions (present or absent), prior MS treatment and age (<40, ≥ 40 years).
- Patients with an initial disability progression during the blinded treatment period who discontinue the treatment early and do not have a subsequent visit with EDSS measurement are imputed as a CDP event or censored with a random probability of 50%.
- Patients with an initial disability progression during the blinded treatment period who discontinue the treatment early and do not have a subsequent visit with EDSS measurement are imputed as having a CDP event.

Source: CSR

12-Week Confirmed Disability Improvement

The results of 12-Week confirmed disability improvement are discussed in the section under Efficacy Results from Reviewer's Analyses

MRI Endpoints

Results from analyses of MRI endpoints are presented under Efficacy Results from Reviewer's Analysis as incorrect counts of patients and scans were presented by the sponsor.

3.2.1.6 Efficacy Results from Reviewer's Analyses

In WA21092, one patient in the IFN group had missing baseline EDSS score and was not included in primary and secondary analyses. The patient discontinued study at Day 18 and did not have relapse. The exclusion of this patient was not indicated in the sponsor's report.

3.2.1.6.1 Results from Analysis of Primary Endpoint ARR

Results from analyses of ARR reported by the sponsor, including sensitivity analyses, were confirmed with the exception of sensitivity analyses with 50% and 100% imputation for missing observations, which were not performed by the reviewer. The sensitivity analyses with 50% and 100% imputation for missing observations were considered biased in favor of ocrelizumab due to the higher rate of discontinuation in the placebo group compared with ocrelizumab group.

The two studies had similar rates of confirmed relapse and rate ratios. These results appeared to be consistent across various sensitivity analyses.

3.2.1.6.2 Analysis of 12-Week CDP

For results of 12-week disability progression, the sponsor performed sensitivity analysis by imputing the event data for discontinued patients, which was specified in SAP. Patients who had initial onset but did not have progression confirmed before discontinuation were considered of having the progression confirmed. Since IFN group had higher discontinuation rate, such imputation methods could be biased in favor of OCR group and would not address the possible impact of discontinuation on disability progression outcomes.

An additional sensitivity analysis was performed by the reviewer. Instead of assuming all initial disability progressions were confirmed for patients who discontinued, an opposite assumption was made that no initial disability progression was confirmation after discontinuation, i.e., patients who discontinued before confirmation could be made were censored at discontinuation. Results from this sensitivity analysis yielded slightly higher hazard ratios than the ones from the

primary analysis. No large discrepancies from the primary analysis were found under this sensitivity analysis. The following table presents results.

Table 10 Assessment of Effect of Discontinuation on Efficacy Results

	WA21092		WA21093	
	IFN beta-1a N=410	OCR 600 mg N=410	IFN beta-1a N=418	OCR 600 mg N=417
Number (%) discontinued	70 (17.1%)	44 (10.7%)	98 ((23.4%)	57 (13.7%)
Mean duration (days)	229	188	209	227
No CDP confirmation after discontinuation				
p-value (logrank)		0.0195		0.0511
HR		0.589		0.681

Source: Reviewer's analysis

3.2.1.6.3 Analysis of Other Secondary Endpoints

Confirmed Disability Improvement

There was a discrepancy between the two studies in the results of 12-Week CDI. A statistically significant treatment difference with a p-value of 0.0106 was found in Study WA21092 while Study WA21093 showed little difference between the two treatment groups in 12-Week CDI. The following table presents the results.

Table 11 Summary of Analysis of 12-Week CDI by Study

	WA21092		WA21093	
	IFN beta-1a N=410	OCR 600 mg N=410	IFN beta-1a N=418	OCR 600 mg N=417
Number (5) with baseline EDSS $\geq$ 2	306 (74.6%)	310 (75.6%)	308 (73.7%)	318 (76.3%)
Number (%) of 12-week CDI	38 (12.4%)	62 (20.0%)	58 (18.8%)	68 (21.4%)
p-value (CMH)		0.0106		0.4019
Number with both CDP and CDI	2	1	0	2

Source: Reviewer's analysis

A few issues were identified with how disability improvement was defined in the protocol. First, it only applied to a subset of the population with a baseline EDSS score of at least 2 and about a quarter of the patients in both studies were not included in this analysis. Second, it did not distinguish between the short term fluctuation and long term benefit. Patient could have an improvement with a progression later or vice versa as shown in Table 11. Lastly, it did not require that the EDSS score at the end of the study to be no higher than baseline EDSS score to insure that the improvement was maintained. The endpoint of disability improvement needed to be well defined in order to be considered meaningful, in the reviewer's opinion.

MRI Endpoints

All three MRI endpoints showed statistically significant treatment differences. The largest treatment difference occurred in total number of Gd-enhancing lesions; the adjusted mean number of lesions per scan was 0.286 for the IFN beta-1a group and 0.016 for the ocrelizumab group with a rate ratio of 0.058 ($p < 0.0001$). The results from analysis of MRI endpoints are presented in the following table. These results are slightly different from the ones obtained by the sponsor mainly in the counts of patients and scans. Some patients did not have baseline observations for the MRI endpoint in the analysis and could not be included in the analysis that used baseline as a covariate.

Table 12 Summary of Results from Analysis of MRI Endpoints

	WA21092		WA21093	
	IFN beta-1a (N=411)	OCR 600 mg (N=410)	IFN beta-1a (N=418)	OCR 600 mg (N=417)
N	374	384	375	385
Total number of T1 lesion	334	21	465	21
Total number of scans	1058	1109	1017	1105
Unadjusted lesions per scan	0.316	0.019	0.457	0.019
Adjusted rate	0.286	0.016	0.416	0.021
Adjusted rate ratio		0.058		0.051
p-value		<0.0001		<0.0001
N	375	388	376	387
Number of T2 new/enlarging	1913	429	2103	380
Total number of scans	1060	1119	1025	1115
Unadjusted lesions per scan	1.805	0.383	2.052	0.341
Adjusted rate	1.413	0.323	1.904	0.326
Adjusted rate ratio		0.229		0.171
p-value		<0.0001		<0.0001
N	374	384	375	385
Number of T1 hypo	1306	564	1484	567
Total number of scans	1058	1109	1016	1106
Unadjusted lesions per scan	1.234	0.509	1.461	0.513
Adjusted rate	0.982	0.420	1.255	0.449
Adjusted rate ratio		0.428		0.357
p-value		<0.0001		<0.0001

Source: Review's analysis

3.2.1.6.4 Summary of Reviewer's Analysis

Overall, the two studies WA21092 and WA21093 showed generally similar and consistent treatment effects in reducing relapse rate and delaying disability progression. Statistically significant treatment differences were also found in MRI endpoints. The overall missing scans were 14% to 19% for IFN beta-1a group and were around 10% for ocrelizumab group.

3.2.2 Evaluation of Efficacy for Protocol WA25046

3.2.2.1 Study Design

The primary objective of this study was to investigate the efficacy of ocrelizumab compared with placebo in patients with **PPMS**. The primary endpoint was the time to onset of CDP over the treatment period, defined as an increase in the EDSS score that was sustained for at least 12 weeks.

Study WA25046 was a Phase III, randomized, double-blind, parallel-group, multicenter study to evaluate the efficacy and safety of 600 mg ocrelizumab, administered as two 300 mg IV infusions separated by 14 days (subsequently referred to as one treatment dose) at a scheduled interval of every 24 weeks, as compared with placebo in adults with PPMS.

A schematic of the study design is presented in Figure 5.

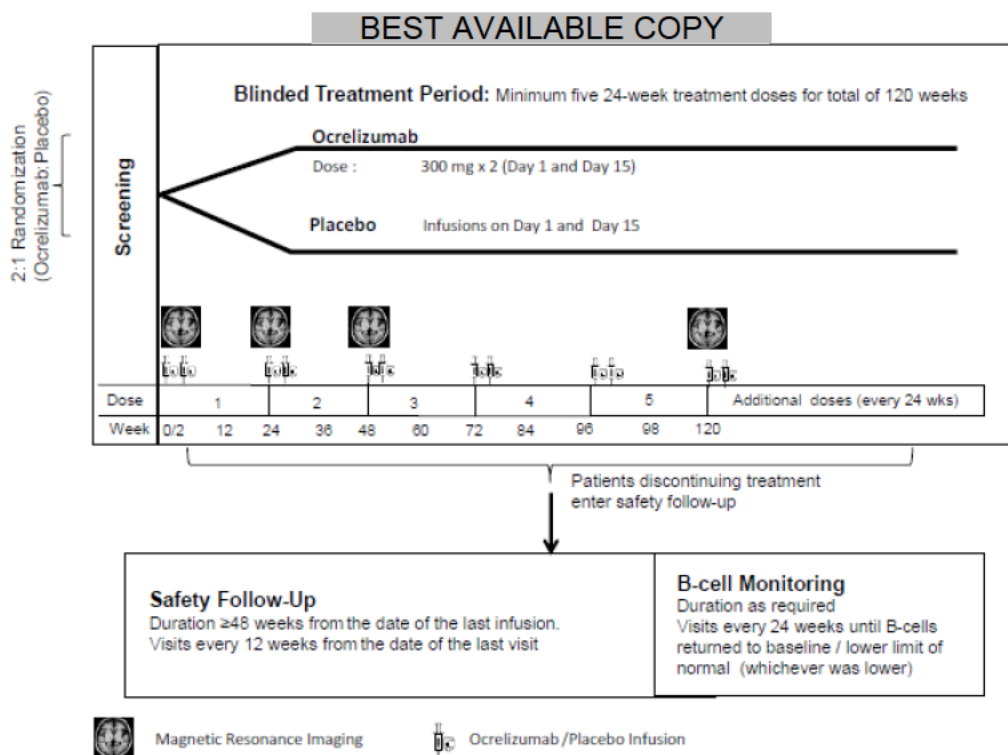


Figure 5 Overview of the Study Design

Source: CSR

The study consisted of a screening period, a double-blind treatment period, and a safety follow-up (SFU) followed by B-cell monitoring, during which patients did not receive any study treatment.

Eligible patients were randomized 2:1 to a regimen of either 600 mg ocrelizumab or placebo, respectively. Patients were treated for a variable duration with the clinical cut-off date (CCOD) occurring after a minimum of 120 weeks and approximately 253 events had been accrued, resulting in at least five treatment doses per patient. Until that time, patients were to continue with blinded dosing every 24 weeks.

Patients who prematurely discontinued study medication during the double-blind treatment period attended a withdrawal visit for the purpose of collecting end-of-treatment assessments.

The study enrolled PPMS patients aged 18 to 55 years with an EDSS score of 3.0-6.5 and duration of disease less than 10 years if EDSS ≤ 5 or less than 15 years if EDSS > 5 at screening.

3.2.2.2 Study Endpoint

The primary efficacy endpoint was the time to onset of 12-week CDP during the double-blind treatment period.

Disability progression was defined as an increase in EDSS from baseline of ≥ 1.0 point, if the baseline EDSS value was ≤ 5.5 points, or ≥ 0.5 points, if the baseline EDSS was > 5.5 points. Assessments within 30 days after a protocol-defined relapse were not used for confirmation of disability progression. The initial disability progression had to occur when the patient was still on treatment. Non-confirmatory EDSS assessments (if any) between the initial and confirmation of disability progression had to fulfill the requirements for progression.

Secondary Endpoints were rank ordered as follows:

- Time to onset of CDP for at least 24 weeks (≥ 161 days, 24-week CDP)
- Change in T25-FW from baseline to Week 120
- Change in total volume of T2 lesions from baseline to Week 120
- Percent change in total brain volume from Week 24 to Week 120
- Change in PCS (SF-36) from baseline to Week 120

3.2.2.3 Statistical Analysis Methods

3.2.2.3.1 General Considerations

All efficacy analyses were to be performed using the intent-to-treat (ITT) population, which was defined as all randomized patients.

For the primary endpoint and secondary efficacy endpoints related to confirmed disability progression, where the initial event of worsening occurred during the double-blind treatment period, confirmation of the event could utilize EDSS scores that were collected during SFU. All other endpoints were to use only data collected from patients during the double-blind treatment period in accordance with the SAP.

All primary and secondary efficacy endpoints were to be analyzed, adjusted or stratified by geographical region (United States vs. ROW) and age (≤ 45 years versus > 45). For CDP, if both screening EDSS and randomization EDSS scores were collected, the average of both values was to be considered as baseline EDSS score for all analyses. If one of the EDSS scores was missing, the other was to be used. For all other assessments, the baseline value was to be used as the last non-missing value on or before the date of the first infusion of study drug.

The null hypothesis was that there was no difference in the underlying endpoint between the ocrelizumab and placebo groups. All statistical hypotheses for the primary and secondary endpoints and treatment comparisons were to be tested at the 5% significance level against two-sided alternatives. Following the analysis of primary endpoint, secondary endpoints were to be analyzed in a hierarchical order in the given rank order.

3.2.2.3.2 Analysis of the Primary Endpoint

The time to CDP for the ocrelizumab and the placebo groups was to be compared using a two-sided log-rank test stratifying by geographic region (US versus ROW) and age (≤ 45 versus > 45). The proportion of patients with confirmed disability progression was to be estimated using Kaplan-Meier methodology. The overall hazard ratio was to be estimated using a stratified Cox regression model with the same stratification factors used in the log-rank test.

The protocol stated that a PPMS patient who experienced initial disease progression (IDP) had an increased risk of disability progression compared to other patients without an initial event who were still ongoing in the treatment period. Patients who had an IDP and then discontinued the treatment early with no confirmatory EDSS assessments were, therefore, not censored as this would introduce substantial bias. This IDP was used as an event and these events were subsequently referred to as imputed events.

Patients who had initial disability progression with no confirmatory EDSS assessment and who were on treatment at time of CCOD were censored at the date of their last EDSS assessment. Patients who did not have initial disability progression at time of CCOD, time of early discontinuation, or loss to follow up were censored at the date of their last EDSS assessment that occurred during the treatment period.

The CCOD occurred when all patients had completed at least 120 weeks on study and the number of CDP events for the primary analysis was approximately 253.

Sensitivity Analyses of Primary Endpoint

Pre-specified sensitivity analyses of the primary endpoint of time to onset of CDP for at least 12 weeks were conducted as follows:

- PP population

- ITT population with multiple imputation (patients with initial disability progression who discontinued treatment early with no confirmatory EDSS assessment were imputed as having CDP with 50% probability)
- ITT population with censoring of the imputed events (patients with initial disability progression who discontinued treatment early with no confirmatory EDSS assessment were censored as not having CDP)
- Influence of early progression events, by omission of EDSS assessments from randomization to Week 12 (≤ 83 days after randomization) (ITT Population)
- Analysis using the planned number of patients (used for study sample size calculation), i.e., using the first 630 patients randomized (ITT Population)
- Adjustment with additional strata for baseline presence of T1 Gd-enhancing lesions (present / absent) and baseline EDSS (≤ 5.5 vs. > 5.5)
- Exclusion of patients with clinical relapses (including protocol-defined relapses)

3.2.2.3.3 Analysis of Secondary Endpoints

Each secondary efficacy endpoint was tested in the given hierarchical order, only if the primary endpoint and each preceding endpoint had reached the significance level of 0.05. The statistical methods employed for each type of endpoint are summarized in Table 13.

APPEARS THIS WAY ON ORIGINAL

Table 13 Statistical Analysis of Primary and Secondary Efficacy Endpoints

Endpoint	Statistical Model	Stratification/adjusting factors
Primary Endpoint		
Time to onset of CDP for at least 12 weeks	Log-rank test for p-value, Cox regression (for estimation of hazard ratio (HR))	Age (≤ 45 vs. > 45 years), geographical region (US vs. ROW)
Secondary Endpoints		
Disability		
Time to onset of CDP for at least 24 weeks	Log-rank test for p-value, Cox regression (for estimation of hazard ratio (HR))	Age (≤ 45 vs. > 45 years), geographical region (US vs. ROW)
Change in Timed 25-Foot Walk Relative Ratio to Baseline at Week 120	ranked ANCOVA with LOCF for p-value; MMRM for treatment estimates	Baseline T25FTW, age (≤ 45 vs. > 45 years), geographical region (US vs. ROW)
Brain MRI		
T2 Lesion Volume Relative Ratio to Baseline at Week 120	ranked ANCOVA with LOCF for p-value; MMRM for treatment estimates	Baseline T2 lesion volume, age (≤ 45 vs. > 45 years), geographical region (US vs. ROW)
Percent Change from Week 24 to Week 120 in Total Brain Volume	MMRM	Week 24 brain volume age (≤ 45 vs. > 45 years) geographical region (US vs. ROW)
Quality of Life		
Mean change from baseline in SF-36 PCS	MMRM	Baseline SF-36 PCS, age (≤ 45 vs. > 45 years), geographical region (US vs. ROW)

Source: CSR

Sensitivity Analyses of Secondary Endpoint - Time to Onset of CDP for at least 24 weeks

Pre-specified sensitivity analyses of the secondary endpoint of time to onset of CDP for at least 24 weeks were conducted as follows:

- Influence of early progression events, by omission of EDSS assessments from randomization to Week 12 (≤ 83 days after randomization) (ITT Population)
- Exclusion of patients with clinical relapses (including protocol-defined relapses)

3.2.2.4 Patient Disposition, Demographic and Baseline Characteristics

3.2.2.4.1 Patient Disposition

A total of 943 patients were screened and 732 patients were randomized into the study, of which 725 received at least one dose of placebo or ocrelizumab. Patients were enrolled at 182 investigational sites across 29 countries.

Patients were randomized in a ratio of 2:1 (ocrelizumab: placebo), stratified by age (≤ 45 vs. > 45 years) and region (US vs ROW). One patient assigned to the placebo group and 6 patients assigned to the ocrelizumab group withdrew before receiving treatment (Figure 6). A total of 549 patients (placebo 162 patients, 66%, vs. OCR 387 patients, 79%) were ongoing with double-blind treatment at the CCOD.

Based on the ITT population, a total of 183 patients (placebo 82 patients [34%], OCR 101 patients [21%]) had withdrawn at the time of the CCOD (including those withdrawn before dosing).

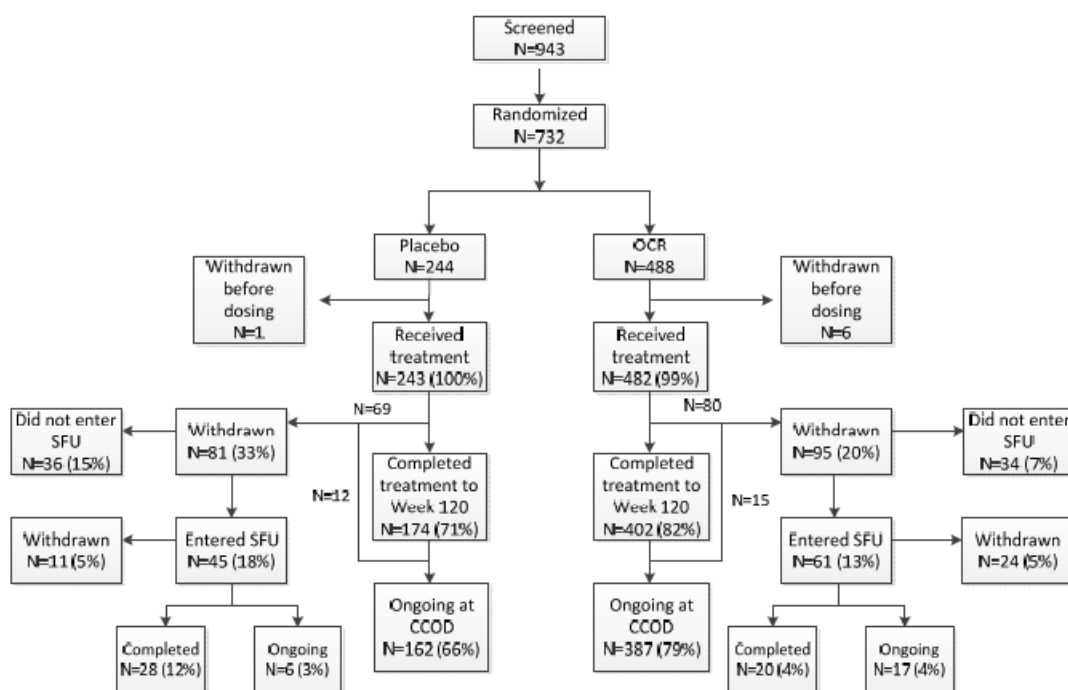


Figure 6 Overview of Patient Disposition

Source: CSR

A higher proportion of patients in the placebo group (34%) withdrew prematurely from treatment during the double-blind treatment period compared to the OCR group (21%). The difference was mainly due to higher incidences of withdrawals due to lack of efficacy (11% versus 4%) and withdrawal by subject (9% versus 5%) in the placebo group versus the OCR group, respectively.

Table 14 Summary of Reasons for Withdrawal from the Double-blind Treatment Period

	Placebo (N=244)	OCR 600mg (N=488)
Discontinued Treatment	82 (33.6%)	101 (20.7%)
ADVERSE EVENT	12 (4.9%)	18 (3.7%)
DEATH	1 (0.4%)	3 (0.6%)
LACK OF EFFICACY	27 (11.1%)	21 (4.3%)
LOST TO FOLLOW-UP	1 (0.4%)	4 (0.8%)
NON-COMPLIANCE	2 (0.8%)	2 (0.4%)
NON-COMPLIANCE WITH STUDY DRUG	2 (0.8%)	2 (0.4%)
OTHER	13 (5.3%)	20 (4.1%)
PHYSICIAN DECISION	2 (0.8%)	6 (1.2%)
PREGNANCY	1 (0.4%)	1 (0.2%)
PROTOCOL VIOLATION	0	2 (0.4%)
WITHDRAWAL BY SUBJECT	21 (8.6%)	22 (4.5%)

Source: CSR

3.2.2.4.2 Patient Demographics

Demographic characteristics of the ITT population appeared to be balanced between the two groups (Table 15). The majority of the patients were white (>90% in both groups). The median age of the population was 46 years (range 18-56 years). Approximately half of the patients were male in both groups. Patients' median weight was 71 kg in the ocrelizumab group and 72 kg in the placebo group. Patients were mainly from Europe (EU, Switzerland and Norway; approximately 64% across both groups) or North America (USA and Canada), Australia and New Zealand (approximately 20% in each group).

For the two stratification factors, approximately 48% of the patients were in ≤ 45 years age group and approximately 52% of the patients were in the > 45 years age group. US patients constituted 14% of the total ITT patient population.

Table 15 Summary of Demographic Data (ITT Population)

	Placebo N=244	OCR 600 mg N=488
Age (years)		
Mean (Median)	44.4 (46.0)	44.7 (46.0)
Age Group, n (%)		
≤ 45	118 (48.4)	230 (47.1)
> 45	126 (51.6)	258 (52.9)
Sex (n (%))		
Male	120 (49.2)	251 (51.4)
Female	124 (50.8)	237 (48.6)
Race (n (%))		
Black or African American	5 (2.0)	9 (1.8)
White	235 (96.3)	454 (93.0)
Other	4 (1.6)	25 (5.1)

Weight (kg) Mean (Median)	72.8 (72.0)	72.5 (71.0)
Region (n (%)) ROW USA	210 (86.1) 34 (13.9)	421 (86.3) 67 (13.7)

Source: CSR

3.2.2.4.3 Patient Baseline Disease Characteristics

The sponsor reported that baseline disease characteristics for MS were similar across the treatment groups (Table 16). The median duration of disease in terms of time from symptom onset was almost 6 years in both groups with a median time since diagnosis of 1.3 years (placebo) and 1.6 years (OCR). The majority of patients (placebo 88% vs OCR 89%) had not received any MS disease-modifying treatment prior to baseline in the previous 2 years.

The mean EDSS score at baseline was 4.7 for both groups (Table 16). Baseline MRI assessments showed that the majority of patients had no T1 Gd-enhancing lesions (placebo 75% vs OCR 73%). The volume and number of T2 lesions as well as normalized brain volume were similar between the groups.

Table 16 Summary of baseline disease characteristics (ITT Population)

	Placebo N=244	OCR 600 mg N=488
Duration since MS symptom onset (years) Mean (Median)	6.1 (5.5)	6.7 (6.0)
Duration since PPMS diagnosis (years) Mean (Median)	2.8 (1.3)	2.9 (1.6)
Prior treatment of MS therapy (N (%)) Yes No	30 (12.3) 214 (87.7)	55 (11.3) 433 (88.7)
Baseline EDSS score Mean (Median) Baseline EDSS category (n (%)) <= 5.5 > 5.5	4.7 (4.5) 163 (66.8) 81 (33.2)	4.7 (4.5) 348 (71.5) 139 (28.5)
Baseline Gd-enhancing T1 lesions Mean (Median) Min - Max Number (%) without lesions	0.6 (0.0) 0.0 - 10.0 183 (75.3)	1.2 (0.0) 0.0 - 77.0 351 (72.5)
Volume of T2 lesions Mean (Median)	10.9 (6.2)	12.7 (7.3)
Normalized brain volume (cm ³) Mean (Median)	1470 (1465)	1463 (1462)

Source: CSR

3.2.2.5 Efficacy Results of Study WA25046 Reported by the Sponsor

3.2.2.5.1 Analysis of the Primary Endpoint – Time to 12-Week CDP

The sponsor reported that treatment with ocrelizumab led to a 24% reduction in the risk of 12-week CDP compared with placebo (Table 17). The Kaplan-Meier survival curves (Figure 7) for time to onset of 12-week CDP show separation from 12 weeks, with a lower proportion of patients in the ocrelizumab group with CDP throughout the treatment period.

Table 17 Time to Onset of 12-week CDP during the Treatment Period with Imputation

	Placebo (N=244)	OCR 600mg (N=488)
Patients included in analysis	244 (100.0%)	487 (100.0%)
Patients with event (%)	96 (39.3%)	160 (32.9%)
Patients without event (%)	148 (60.7%)	327 (67.1%)
Time to event (weeks) Range	0* to 216*	0* to 217*
Stratified Analysis p-value (log-rank)	0.0321	
Hazard Ratio 95% CI	0.76 (0.59, 0.98)	
Time Point Analysis		
48 Weeks		
Patients at risk	189	414
Event rate (%)	16.96	11.05
95% CI	(12.17, 21.76)	(8.22, 13.88)
96 Weeks		
Patients at risk	153	338
Event rate (%)	28.28	24.75
95% CI	(22.43, 34.13)	(20.80, 28.70)
120 Weeks		
Patients at risk	136	304
Event rate (%)	33.98	30.23
95% CI	(27.77, 40.18)	(26.00, 34.45)

Time to CDP (weeks) - Censoring: Event (1=censored, 0=event)

<* Censored observation>

Stratified by Geographical Region (US vs. ROW) and Age (<= 45, >45 years).

Hazard ratios were estimated by stratified Cox regression.

Patient with missing baseline EDSS excluded from analysis. Patients with an initial disability progression during the blinded treatment period who discontinue the treatment early and do not have a subsequent visit with EDSS measurement are imputed as having a CDP event.

Source: CSR

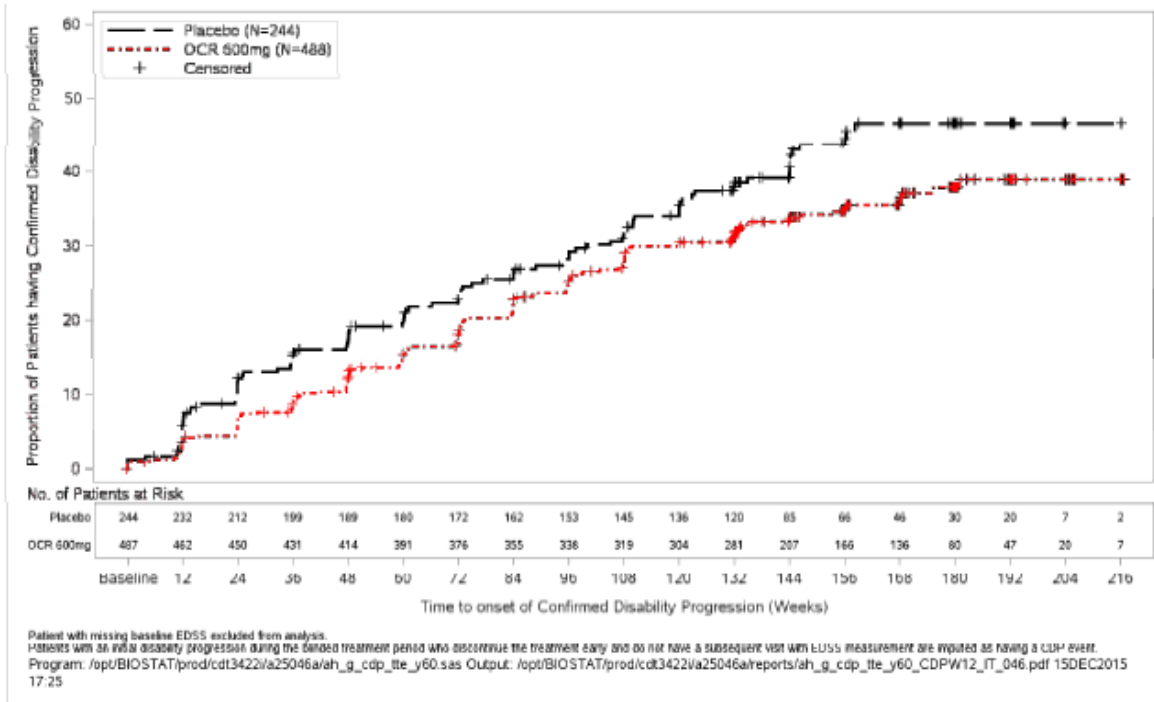


Figure 7 Kaplan-Meier Plot of Time to Onset of Confirmed Disability Progression for At Least 12 Weeks during the Double-blind Treatment Period

Source: CSR

Various sensitivity analyses were performed by the sponsor and the results are presented in Table 18. With regard to imputation of initial disability progression events for patients with discontinuation, the approach of censoring patients at discontinuation for those who had initial disability onset without confirmation resulted in a reduced treatment effect (HR 0.82 [95% CI: 0.63, 1.07], $p=0.1477$).

The sensitivity analysis including progression events after treatment discontinuation resulted in a reduced treatment effect (HR 0.80 [95% CI: 0.62, 1.02], $p=0.0736$).

The sensitivity analysis removing progression events preceded by a PDR resulted in a similar treatment effect to that seen in the primary analysis (HR 0.78 [95% CI: 0.60, 1.01], $p=0.0561$), suggesting the efficacy of ocrelizumab in delaying disability progression was not due to an effect on relapses.

Table 18 Sensitivity analyses of the primary endpoint

Sensitivity Analysis	Patients with event / N		Hazard Ratio (95% CI)	p-value
	Placebo	OCR		
PP Population (summary and plot)	91 /232	153 /469	0.74 (0.57, 0.96)	0.0239
ITT population with multiple imputation (summary)	#	#	0.78 (0.60, 1.02)	#
ITT population without imputation (summary and plot)	84 /244	151 /487	0.82 (0.63, 1.07)	0.1477
Influence of early progression events, by omission of EDSS assessments from randomization to Week 12 (≤ 83 days after randomization) (ITT Population) (summary)	94 /244	160 /487	0.78 (0.60, 1.00)	0.0500
Analysis using planned number of patients, i.e, using first 630 patients randomized (ITT Population) (summary)	83 /209	145 /420	0.79 (0.60, 1.04)	0.0867
Adjustment with additional strata for presence of T1 Gd-enhancing lesions (present / absent) and baseline EDSS (≤ 5.5 vs. > 5.5) (summary)	96 /244	160 /487	0.76 (0.59, 0.98)	0.0321

Sensitivity Analysis	Patients with event / N		Hazard Ratio (95% CI)	p-value
	Placebo	OCR		
Exclusion of patients with clinical relapses (including protocol-defined relapses) (summary and plot)	77 /204	144 /456	0.74 (0.56, 0.98)	0.0324
*Analysis including progression after treatment discontinuation (summary and listing)	97 /244	166 /487	0.80 (0.62, 1.02)	0.0736
*Imputation by efficacy related reason for withdrawal / withdrawal by subject (summary)	92 /244	156 /487	0.77 (0.60, 1.00)	0.0490
*Analysis excluding CDP events where a PDR was experienced from 30 days preceding an IDP or between IDP and CDP (summary)	92 /244	157 /487	0.78 (0.60, 1.01)	0.0561

No p-value calculated for multiple imputation results (50% of events are randomly imputed across groups, therefore the number of events in each group differs with each imputation).

* Post-hoc sensitivity analyses.

Source: CSR

3.2.2.5.2 Analysis of Secondary Endpoints

Time to 24-Week CDP

The sponsor reported that consistent with the primary endpoint, treatment with ocrelizumab led to a 25% reduction in the risk of 24-week CDP compared with placebo (hazard ratio 0.75 [95% CI: 0.58, 0.98], p = 0.0365) (Table 19).

Table 19 Time to Onset of 24-week CDP Weeks during the Treatment Period with Imputation

	Placebo (N=244)	OCR 600mg (N=488)
Patients included in analysis	244 (100.0%)	487 (100.0%)
Patients with event (%)	87 (35.7%)	144 (29.6%)
Patients without event (%)	157 (64.3%)	343 (70.4%)
Time to event (weeks)		
Range	0* to 216*	0* to 217*
Stratified Analysis		
p-value (log-rank)	0.0365	
Hazard Ratio	0.75	
95% CI	(0.58, 0.98)	
Time Point Analysis		
48 Weeks		
Patients at risk	193	421
Event rate (%)	15.26	9.56
95% CI	(10.67, 19.86)	(6.91, 12.22)
96 Weeks		
Patients at risk	157	349
Event rate (%)	26.56	22.38
95% CI	(20.83, 32.29)	(18.56, 26.19)
120 Weeks		
Patients at risk	139	313
Event rate (%)	32.71	28.30
95% CI	(26.56, 38.86)	(24.15, 32.45)
Time to CDP (weeks) - Censoring: Event (1=censored, 0=event)		
<* Censored observation>		
Stratified by Geographical Region (US vs. ROW) and Age (<= 45, >45 years).		
Hazard ratios were estimated by stratified Cox regression.		
Patient with missing baseline EDSS excluded from analysis. Patients with an initial disability progression during the blinded treatment period who discontinue the treatment early and do not have a subsequent visit with EDSS measurement are imputed as having a CDP event.		

Source: CSR

Timed 25-foot Walk Test

The sponsor reported that ocrelizumab treatment resulted in a 29% relative reduction in the percent progression in T25FW over 120 weeks (p = 0.0404) compared to placebo (Table 20).

Table 20 Summary of Analysis of T25FW

	Placebo (N=244)	OCR 600 mg (N=488)
Baseline		
n	239	473
Mean (sec) (SE)	12.781 (1.00)	14.573 (0.95)
Week 120 (Ratio Relative to Baseline)		
n	174	397
Adjusted Geometric Mean (% change)	1.551 (55.10)	1.389 (39.93)
95% CI Adjusted Geometric Mean	1.399, 1.720	1.292, 1.494
Ratio of Adjusted Geometric Means		0.896
95% CI for Ratio of Adjusted Geometric Means		0.792, 1.013
Relative Reduction (%)		29.337
95% CI for Relative Reduction		-1.618, 51.456
p-value (Ranked ANCOVA)		0.0404

Source: [summary](#) and [relative reduction](#)

Relative reduction was calculated as -Relative change = - (OCR response-Placebo response)/Placebo response*100%. The 95% CI for relative reduction was obtained using the Bootstrap method.

Source: CSR

Change in the Total Volume of T2 Lesions from Baseline to Week 120

The sponsor reported that Ocrelizumab decreased the percentage change in total volume of T2 hyperintense lesions from baseline to Week 120 (decrease of 3.4%) compared with an increase for patients on placebo (increase of 7.4%; $p < 0.0001$). Table 21 presents the results.

Table 21 Ratio and Percentage Change in Total T2 Lesion Volume Relative to Baseline

	Placebo (N=244)	OCR 600 mg (N=488)
Baseline		
n	234	464
Mean (cm ³) (SE)	11.039 (0.858)	12.761 (0.709)
Week 120 (Ratio Relative to Baseline)		
n	183	400
Adjusted Geometric Mean	1.074	0.966
95% CI Adjusted Geometric Mean	1.050, 1.099	0.950, 0.983
Ratio of Adjusted Geometric Means		0.900
95% CI for Ratio of Adjusted Geometric Means		0.876, 0.924
Adjusted Geometric Mean (% change)	7.426	-3.366
95% CI for Adjusted Geometric Mean (% change)	4.967, 9.942	-4.987,-1.718
p-value (Ranked ANCOVA)		<0.0001

Source: CSR

Other secondary Efficacy Endpoints

Other secondary endpoints (percent change in total brain volume from Week 24 to Week 120 and change in PCS (SF-36) from baseline to Week 120) were not considered clinically validated by the Division and results are not presented.

3.2.2.6 Reviewer's Analysis

3.2.2.6.1 Analysis of the Primary Endpoint

The results from the analyses reported by the sponsor were confirmed. There are 2 issues in the primary analysis that need to be pointed out.

First, the primary analysis was based on the imputed events of 12-week CDP. It was specified in the protocol and SAP that if a patient had initial onset of CDP then withdrew from the study, the CDP would be considered, or imputed, as confirmed. There were 21 such events, 12 in the placebo group and 9 in the ocrelizumab group. The placebo group had higher rate of CDP as well as high rate of discontinuation compared to ocrelizumab group, and such imputed CDP represented 12.5% and 5.6% of the total CDP events in the placebo group and ocrelizumab group, respectively.

Table 22 Confirmed Disability Progression without confirmation due to withdrawal

	Placebo N=244	Ocrelizumab N=487
Number of patients with 12-week CDP		
With imputation	96 (39.3%)	160 (32.9%)
Without imputation	84 (34.4%)	151 (31.0%)
Total number (%) withdrew	82 (33.6%)	101 (20.7%)
# progression without confirmation due to WD*	12	9
% of such progression of total withdrew	14.6%	8.9%
% of such progression of total progression	12.5%	5.6%

* Withdrawal

Source: Reviewer's analysis

In a pre-specified sensitivity analysis based on the un-imputed events, it was reported by the sponsor and confirmed by the reviewer that the treatment difference was reduced with a hazard ratio of 0.82 and a p-value of 0.1477 (Table 18). There were similar findings in the sensitivity analysis of 24-week CDP where a reduced treatment difference with hazard ratio of 0.82 (p=0.1884) was observed. The results from these sensitivity analyses casted a question of whether the withdrawal had contributed to the size and significance of the treatment difference.

Second, it appears from Figure 7 that the two curves of the Kaplan-Meier plot for the time to CDP separated at around Week 12 and kept near parallel through the end of the treatment period. In order to exam whether treatment of ocrelizumab showed any effect after Week 12, analyses were performed on the patients who were at the risk at the beginning of weeks 24, 48, 72, 96, and 120. For example, in order to see if there was any treatment effect left after week 24, patients who had 12-week CDP prior to Week 24 or withdrew by Week 24 were removed from analysis and the same Cox model was applied to the remaining 660 patients (211 in the placebo group and 449 in the OCR group) who were still at risk at week 24 assessment. The analysis showed that the treatment difference was reduced as indicated by an increase in hazard ratio from 0.759 to 0.858 (Table 23).

The procedure was repeated at weeks 48, 72, 96, and 120, with fewer patients remaining at risk at each assessment followed. The following table presents the number of patients remaining at risk at each time point, the number of CDP events from the reduced population from the time point to the end of the study, and the hazard ratio estimated from the reduced patient population.

Table 23 Analysis of 12-week CDP in the Patient Still at Risk at Weeks 24, 48, 72, 96, and 120

	Placebo N=244	Ocrelizumab N=487
Beginning of the Interval = Day 1 (primary)		
Number at the risk	244	487
Number (%) of CDP from Day 1	96 (39.3%)	160 (32.9%)
HR		0.759
Beginning of the Interval = Week 24		
Number at the risk	211	449
Number (%) of CDP from week 24	72 (34.1%)	138 (30.7%)
HR		0.858

Beginning of the Interval = Week 48 Number at the risk Number (%) of 12-week CDP HR	188 56 (29.8%)	414 108 (26.1%) 0.852
Beginning of the Interval = Week 72 Number at the risk Number (%) of 12-week CDP HR	172 44 (25.6%)	375 81 (21.6%) 0.806
Beginning of the Interval = Week 96 Number at the risk Number (%) of 12-week CDP HR	153 31 (20.3%)	334 46 (13.8%) 0.634
Beginning of the Interval = Week 120 Number at the risk Number (%) of 12-week CDP HR	136 19 (14.0%)	303 22 (7.3%) 0.464

Source: Reviewer's analysis

Although at each of the time points OCR group had smaller proportion of patients with 12-week CDP, it was clear that after Week 24 the treatment difference was reduced with higher hazard ratio until Week 96. The low hazard ratio at Week 120 was probably due to the imputed CDP of 6 patients in the placebo group versus one patient in the OCR group after week 120.

3.2.2.6.2 Analysis of Secondary Endpoints

Time to 24-week Confirmed Disability Progression

The results from analysis of 24-week CDP were almost identical to the ones from analysis of 12-week CDP with the same issues.

Analyses of Timed 25-Foot Walk Score

There was a substantive revision of SAP (Version E) made after the study completion and data cut-off date (July 24, 2015). The Version E of the SAP was submitted to the Division on September 4, 2015 prior to data unblinding. Because the submission came in at such a late stage, the Division did not have sufficient time to review the revised SAP and did not provide comments to the sponsor.

The changes made in the Version E of the SAP involved analyses of secondary endpoints of timed 25-foot walk (T25FW) and T2 lesion volume; both were revised previously in Version D.

The analysis of T25FW was changed from MMRM model on the absolute change from baseline to Week 120 to ranked ANOVA on the last observation carry forward (LOCF) in percentage change from baseline to Week 120. The treatment effect was to be derived from the MMRM analysis on the log transformed ratio of post-baseline/baseline value.

The sponsor stated that the rationale for the change was that the experience of trials WA21092 and WA21093 and blinded review of data from this trial indicated that the normal assumption would be violated in the distribution of the residuals.

In the study report, the sponsor reported a p-value of 0.0404 from the ranked ANOVA analysis for T25FW and the treatment difference was represented by a ratio of adjusted genomic mean of 0.896 from the MMRM analysis on the log transformed ratio of T25FW at each visit and T25FW at baseline ($\log[T25FW@visit\ x/T25FW@baseline]$).

The MMRM analysis performed by the reviewer resulted in a same adjusted genomic mean of 0.896 with a p-value of 0.0783 from the MMRM analysis.

The sponsor reported a p-value of 0.0404 from the rank ANOVA analysis. A total of 20 patients (5 in the placebo group and 15 in the ocrelizumab group) did not have any post-baseline assessment scores and their baseline score was carried forward (specified in the revised SAP) in the rank ANOVA analysis.

Given that about 70% of the patients in both groups had an increase in T25FW walking time, assigning a 0 change to 5 patients in the placebo group and 15 patients in the ocrelizumab group gave ocrelizumab a benefit in the analysis of treatment difference. The reviewer performed the same analysis without carrying forward the baseline value (i.e., patients without post-baseline scores were excluded) and a p-value of 0.0528 was obtained.

The following table presents a summary of descriptive statistics and the results from various analyses performed by the reviewer. These results appeared to be consistent regardless of which model was used with marginally insignificant treatment difference at best, except the one from the ranked ANOVA analysis performed by the sponsor. Moreover, larger increase in the walking speed in the ocrelizumab group compared to placebo group was observed from Week 48 to Week 96.

Table 24 Descriptive statistics and sensitivity analyses of T25FW

T25FW Scores (seconds)	Placebo N=244	Ocrelizumab N=487
Patients with both baseline and post-baseline assessment scores	N=239	N=473
Baseline		
Mean (SD)	12.78 (15.48)	14.57 (20.55)
Median	7.25	7.75
LOCF at Week 120		
Mean (SD)	28.61 (49.37)	25.91 (42.17)
Median	9.55	9.79
Change from baseline at Week 120 (LOCF)		
Mean (SD)	15.83 (43.33)	11.33 (38.38)
Median	1.25	0.85
Change from baseline to Week 48 (OC)		
N	218	450
Mean (SD)	7.76 (30.56)	2.59 (22.60)
Median	0.55	0.20

Change from week 48 to Week 96 (OC)		
N	187	417
Mean (SD)	4.49 (29.60)	5.26 (26.95)
Median	0.35	0.45
Change from week 96 to Week 120 (OC)		
N	171	389
Mean (SD)	4.58 (26.82)	2.99 (24.67)
Median	0.25	0.15
Primary analysis for p-value (rank ANOVA on percentage change from baseline LOCF)		P=0.0528
Primary analysis for treatment difference		
Adjusted genomic mean from MMRM		0.896
p-value		0.0783
Rank ANOVA on change from baseline LOCF		P=0.0743

Source: Reviewer's analysis

Analysis of Change from Baseline to Week 120 in T2 Lesion Volume

The analysis of T2 lesion volume was changed from MMRM on percentage change from baseline to Week 120 to ranked ANCOVA on absolute change from baseline in Version E of the SAP.

There were 3 MRI scans performed during the trial at Weeks 24, 48, and 120. At each scan, the mean T2 lesion volume in the ocrelizumab group was generally decreased from the previous scan and the baseline one for both treatment groups with larger decrease in the ocrelizumab group.

Table 25 Summary of T2 Lesion Volumes at Weeks 24, 48, and 120

	Placebo N=244	Ocrelizumab N=487
Patients with both baseline and post-baseline assessment scores	N=229	N=456
Baseline		
Mean (SD)	11.19 (13.22)	12.67 (15.17)
Median	6.52	7.24
Week 24		
Mean (SD)	11.15 (13.14)	12.39 (14.90)
Median	6.42	7.05
Change from baseline to Week 24		
Mean (SD)	-0.04 (1.05)	-0.28 (1.25)
Median	-0.04	-0.11
Week 48		
N	215	441
Mean (SD)	11.02 (13.26)	12.24 (14.89)
Median	6.38	6.90
Change from Week 24 to Week 48		
Mean (SD)	0.13 (1.04)	-0.13 (0.90)
Median	0.00	-0.02

Week 120		
N	176	396
Mean (SD)	10.84 (12.65)	11.85 (13.32)
Median	5.82	7.20
Change from Week 48 to Week 120		
Mean (SD)	0.58 (1.66)	0.04 (0.81)
Median	0.07	0.00
Change from baseline to Week 120 LOCF		
N	230	461
Mean (SD)	0.53 (2.38)	-0.38 (1.51)
Median	0.05	-0.12
Primary analysis for p-value (rank ANOVA on percentage change from baseline LOCF)		P<0.0001
Primary analysis for treatment difference		
Adjusted genomic mean from MMRM		0.901
p-value		<0.0001
Rank ANOVA on change from baseline LOCF		P<0.0001

Source: Reviewer's analysis

3.2.2.6.3 Other Analyses Performed in Response to the Outcomes from Site Inspection

A few issues, including some serious ones, were identified during the site inspection. Based on the Clinical Inspection Summary written by Dr. Cara Alfaro, the clinical investigator inspections revealed regulatory violations for three of the clinical investigator sites (Huang, Selmaj, and De Seze).

At site 208392 (Selmaj), EDSS assessments for Protocol WA25046 were not entered into the computer tablet at the time the assessments were completed, as specified in the protocol, for thirteen of nineteen randomized subjects. For these EDSS data entries, there were no source documents available to verify data integrity. The inspection of this site has been preliminarily classified as Official Action Indicated (OAI). Dr. Alfaro recommends the review division to perform sensitivity analyses excluding this site.

Site 208392 randomized 19 patients: 5 to placebo group and 14 to ocrelizumab group. Four of the 5 patients randomized to placebo and 7 of the 14 patients randomized to ocrelizumab had 12-week CDP. Excluding the patients from this site resulted in a p-value of 0.0454 from the log-rank test (the primary analysis). A hazard ratio of 0.769 was obtained from the Cox regression model.

At another site (site # 208241, De Seze), the field investigator noted two instances of study drug administration errors. After the infusion at Visit 18, Subject 47109 (randomized to placebo) experienced a skin reaction that had not been experienced in prior infusions. The drug dispensing error was discovered when the investigators requested the pharmacist to check whether there had been a drug dispensing error. Due to this error, Subject 47109 received ocrelizumab instead of placebo at this visit.

A similar drug dispensing error was discovered for Subject 47103 (randomized to ocrelizumab) at Visit 21. This error was noted when kit number 115348 which was to be dispensed for Subject 47117 was noted to be missing.

Dr. Alfaro commented that the study drug administration errors for both subjects were noted in the aex.xpt dataset. Therefore, it appeared that the sponsor was made aware of this drug dispensing error. Due to the potential for unblinding, Dr. Alfaro recommended that the review division perform a sensitivity analysis excluding these two subjects.

It was found that Subject 47109 (randomized to placebo group) had 12-week CDP occurred at Day 1009 and Subject 47103 (randomized to ocrelizumab) did not have 12-week CDP and was censored at Day 1344. An analysis was performed with the exclusion of these two subjects. The results were little changed with a p-value of 0.0384 from the log-rank analysis and a hazard ratio of 0.766 from the Cox regression model.

3.2.2.6.4 Summary of Reviewer's Analysis

Results from Study WA25046 on PPMS patients appear to be indicative of efficacy in delaying the disability progression with the treatment of ocrelizumab. However, the evidence of effectiveness is weak. Of the total 256 CDP events, 21 events were imputed and not actually confirmed. Statistical significance of the treatment difference was lost in the analysis without imputation.

3.3 Evaluation of Safety

Please refer to Evaluation of Safety by Dr. Gerard Boehm.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

For RMS trials WA21092 and WA21093, subgroup analyses of ARR and 12-week CDP by gender and age group are presented in Tables 26 and 27, respectively. Treatment differences between the gender and age groups were generally similar with the exception that the older age group showed smaller treatment effect in ARR.

Table 26 Subgroup Analysis of ARR by Gender and Age, Studies WA21092 and WA21093

	WA21092		WA21093	
	IFN beta-1a N=410	OCR 600 mg N=410	IFN beta-1a N=418	OCR 600 mg N=417
Age Group (years)				
< 40				
N	242	244	241	252
ARR	0.358	0.152	0.362	0.146
Rate Ratio		0.423		0.403
≥ 40				
N	168	166	177	165
ARR	0.245	0.169	0.173	0.214
Rate Ratio		0.692		0.808
Sex				
Female				
N	271	270	280	271
ARR	0.306	0.152	0.165	0.301
Rate Ratio		0.498		0.547
Male				
N	139	140	138	146
ARR	0.267	0.162	0.120	0.231
Rate Ratio		0.607		0.520

Source: Reviewer's analysis

Table 27 Subgroup Analysis of 12-week CDP by Gender and Age, Studies WA21092 and WA21093

	WA21092		WA21093	
	IFN beta-1a N=410	OCR 600 mg N=410	IFN beta-1a N=418	OCR 600 mg N=417
Age Group (years)				
< 40				
N	242	244	241	252
Number (%) of CDP	23 (9.50)	13 (5.33)	30 (12.45)	21 (8.33)
Hazard ratio		0.538		0.637
≥ 40				
N	168	166	177	165
Number (%) of CDP	27 (16.07)	18 (10.84)	33 (18.64)	23 (13.94)
Hazard ratio		0.624		0.606
Sex				
Female				
N	271	270	280	271
Number (%) of CDP	29 (10.70)	18 (6.67)	41 (14.64)	30 (11.07)
Hazard ratio		0.573		0.693
Male				
N	139	140	138	146
Number (%) of CDP	21 (15.11)	13 (9.29)	22 (15.94)	14 (9.59)
Hazard ratio		0.583		0.535

Source: Reviewer's analysis

For the PPMS trials WA25046, subgroup analyses of the 12-week CDP by gender and age group are presented in Tables 28. Ocrelizumab did not show treatment benefit numerically or statistically among female patients, which constituted about half of the patient population.

Table 28 Subgroup Analysis of 12-week CDP by Gender and Age, Study 25046

	Placebo N=244	OCR 600 mg N=487
Age Group (years)		
≤ 45		
N	118	230
Number (%) of CDP	49 (41.5%)	71 (30.9%)
Hazard ratio		0.641
> 45		
N	126	257
Number (%) of CDP	47 (37.3%)	89 (34.6%)
Hazard ratio		0.884
Sex		
Female		
N	124	236
Number (%) of CDP	44 (35.5%)	85 (36.0%)
Hazard ratio		0.944
Male		
N	120	251
Number (%) of CDP	52 (43.3%)	75 (29.9%)
Hazard ratio		0.614

Source: Reviewer's analysis

4.2 Other Special/Subgroup Populations

Region, baseline EDSS score and baseline body mass index (BMI) were considered relevant factors that might affect the efficacy results. Therefore, subgroup analyses of ARR and 12-week CDP by these factors were performed. Results of analyses by these sub-populations in ARR and 12-week CDP for Studies WA21092 and WA21093 are presented in Tables 29 and 30, respectively. Results of analysis by these sub-populations in 12-week CDP for PPMS trial WA25046 are presented in Table 31.

Table 29 Subgroup Analyses of ARR by Region, Baseline EDSS Score and BMI, Studies WA21092 and WA21093

	WA21092		WA21093	
	IFN beta-1a N=411	OCR 600 mg N=410	IFN beta-1a N=418	OCR 600 mg N=417
Region				
ROW				
N	306	305	304	305
ARR	0.283	0.164	0.291	0.137
Rate Ratio		0.579		0.470
USA				
N	104	105	114	112
ARR	0.308	0.134	0.264	0.199
Rate Ratio		0.435		0.754

Baseline EDSS Score				
< 4				
N	318	314	309	315
ARR	0.217	0.104	0.240	0.116
Rate Ratio		0.480		0.485
≥ 4				
N	92	96	109	102
ARR	0.373	0.249	0.342	0.223
Rate Ratio		0.669		0.652
Baseline BMI				
< 25				
N	209	216	204	190
ARR	0.335	0.176	0.269	0.125
Rate Ratio		0.527		0.463
≥ 25				
N	200	192	208	220
ARR	0.251	0.116	0.298	0.180
Rate Ratio		0.463		0.602

Source: Reviewer's analysis

Table 30 Subgroup Analyses of 12-week CDP by Region, Baseline EDSS Score and BMI, Studies WA21092 and WA21093

	Placebo N=244	OCR 600 mg N=488
Region		
ROW		
N	210	420
Number (%) of CDP	84 (40.0%)	145 (34.5%)
Hazard ratio		0.789
USA		
N	34	67
Number (%) of CDP	12 (35.3%)	15 (22.4%)
Hazard ratio		0.553
Baseline EDSS		
< 4		
N	66	130
Number (%) of CDP	24 (36.4%)	33 (25.4%)
Hazard ratio		0.675
≥ 4		
N	178	357
Number (%) of CDP	72 (40.5%)	127 (35.6%)
Hazard ratio		0.783
BMI Category		
< 25		
N	139	289
Number of CDP	57 (41.0%)	91 (31.5%)
Hazard ratio		0.675
≥ 25		
N	103	196
Number of CDP	39 (37.9%)	69 (35.2%)
Hazard ratio		0.893

Source: Reviewer's analysis

Table 31 Subgroup Analyses of 12-week CDP by Region, Baseline EDSS Score and BMI, Studies WA25046

	Placebo N=244	OCR 600 mg N=487
Region		
ROW		
N	210	420
Number (%) of CDP	84 (40.0%)	145 (34.5%)
Hazard ratio		0.789
USA		
N	34	67
Number (%) of CDP	12 (35.3%)	15 (22.4%)
Hazard ratio		0.553
Baseline EDSS		
< 4		
N	66	130
Number (%) of CDP	24 (35.4%)	33 (25.4%)
Hazard ratio		0.675
≥ 4		
N	178	357
Number (%) of CDP	72 (40.5%)	127 (35.6%)
Hazard ratio		0.783
BMI Category		
< 25		
N	139	289
Number (%) of CDP	57 (41.0%)	91 (35.2%)
Hazard ratio		0.675
≥ 25		
N	109	196
Number (%) of CDP	39 (37.9%)	69 (35.2%)
Hazard ratio		0.893

Source: Reviewer's analysis

No large discrepancies in efficacy among the subgroups were found. The high BMI might have a negative effect on the overall efficacy.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

The two pivotal studies WA21092 and WA21093 conducted in relapsing forms of MS patients produced similar and consistent efficacy results including results on the annualized relapsing rate and confirmed disability progression. No major issues were identified in the efficacy results of the two studies.

In study WA25046 conducted on the PPMS patients, a possible bias in the primary analysis was identified and a sensitivity analysis failed to address the issue.

The primary analysis was based on the imputed events such that 21 of the confirmed disability progression events (12 in the placebo group and 9 in the ocrelizumab group) were imputed and not actually confirmed. These 21 events were imputed as CDP due to patient withdrawal after initial disability progression. Because the placebo group had a higher rate of withdrawal, the primary analysis with imputed events might have contributed bias in favor of ocrelizumab group in determining the treatment difference and its significance. Significance of the treatment difference was lost when censoring was applied at withdrawal to these patients.

It should also be noted that in the PPMS study, the power to detect treatment difference was driven by events. It was estimated that about 253 events in total were needed to detect a treatment difference. A total of 256 events, including 21 imputed events, were actually occurred. Therefore, the loss of significance in the sensitivity analysis on events without imputation could be partly due to the removing of bias and partly due to the loss in the number of events.

The study also showed difference in efficacy between males and females. It appeared that there was little or no treatment benefit among female patients, which constituted half of the patient population.

5.2 Collective Evidence

Studies WA21092 and WA21093 met the study objectives by demonstrating statistically significant treatment effects in the primary endpoint of annualized relapse rate and the key secondary endpoint of disability progression. The results appeared to be consistent under sensitivity analyses and across demographic and baseline characteristics in both studies.

It was in an agreement between the sponsor and the Division that the analysis of disability progression was to be based on the pooled data of the two studies due to the difficulty to achieving the required power to detect the treatment difference in individual studies. However, the two studies were able to show a statistically significant treatment effect individually.

In Study WA25046, which was conducted in PPMS patients, a statistically significant treatment difference was achieved in both 12-week confirmed disability progression (the primary endpoint) and 24-week confirmed disability progression (a secondary endpoint) in the protocol-defined primary analysis. The protocol-defined primary analysis allowed initial onset of disability progression for patients who withdrew from the study to be imputed as confirmed without actual confirmation. In both 12-week CDP and 24-week CDP, the study failed to achieve statistically significant treatment effect in a sensitivity analysis on un-imputed data, which is commonly used in the primary analysis for disability endpoint.

5.3 Conclusions and Recommendations

Studies WA21092 and WA21093 provided consistent evidence that the treatment of ocrelizumab was effective as compared to interferon beta-1a in reducing the relapse rate and delaying the disability progression in patients with relapsing forms of MS.

Study WA25046 provided data that were indicative of efficacy in the treatment of ocrelizumab in delaying the disability progression in patients with PPMS. The evidence of the effectiveness was weakened by the failure of the study to withstand an important sensitivity analysis on un-imputed data, which is commonly used as the standard primary data for disability progression endpoint.

APPEARS THIS WAY ON ORIGINAL

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

XIAORONG YAN
10/26/2016

KUN JIN
10/26/2016
I concur with the review.

HSIEN MING J HUNG
10/27/2016