A member of the AstraZeneca Group

Statistical Analysis Plan Approval

Date: *18Oct2018*
To: Study File
From: PPD
Re: Statistical Analysis Plan Approval for Study *CD-IA-MEDI-551-1155*

The Statistical Analysis Plan, Amendment 4, for Study *CD-IA-MEDI-551-1155* has been reviewed and approved.

Name:	PPD	Signature:	DocuSigned by: PPD 18 October 2018
Role:	Statistician	Date:	
Name:	PPD	Signature:	DocuSigned by: PPD 18 October 2018
Role:	Statistical Programmer	Date:	
Name:	PPD	Signature:	DocuSigned by: PPD 19 October 2018
Role:	Clinical Development Lead	Date:	
Name:	PPD	Signature:	DocuSigned by: PPD 18 October 2018
Role:	Interim Clinical Biostatistics Therapeutic Area Head	Date:	
Name:	PPD	Signature:	DocuSigned by: PPD 18 October 2018
Role:	On Behalf of Head of Clinical Biostatistics & Data Management	Date:	

Statistical Analysis Plan

A Double-masked, Placebo-controlled Study with Open-label Period to Evaluate the Efficacy and Safety of MEDI-551 in Adult Subjects with Neuromyelitis Optica and Neuromyelitis Optica Spectrum Disorders

Protocol Number: CD-IA-MEDI-551-1155

List of Abbreviations

Abbreviation or Specialized Term	Definition
AC	Adjudication Committee
ADA	Anti-drug antibodies
AE	Adverse event
AESI	Adverse event of Special Interest
ANCOVA	Analysis of covariance
AQP4	Aquaporin-4
AZA	Azathioprine
AZDD	AstraZeneca Drug Dictionary
BMI	Body mass index
CF	Counting fingers
CI	Confidence interval
C-SSRS	Columbia-Suicide Severity Rating Scale
ECG	Electrocardiogram
EDSS	Kurtzke Expanded Disability Status Scale
FACS	Flow cytometry assay
FSS	Functional Systems Score
GCP	Good clinical practice
Gd	Gadolinium
HCRU	Health-related quality of life
HM	Hand movement
HR	Hazard ratio
HRQoL	Health-related quality of life
ICC	Intraclass correlation coefficient
IDMC	Independent Data Monitoring Committee
Ig	Immunoglobulin
IP	Investigational product
ITT	Intent-to-Treat
IV	Intravenous
IVIG	Intravenous immunoglobulin
KM	Kaplan-Meier
LP	Light perception
LSMEAN	Least squared mean
MedDRA	Medical Dictionary for Regulatory Activities
MCS	Mental component score
MMRM	Mixed effects model repeated measures
MRI	Magnetic resonance imagining

Abbreviation or Specialized Term	Definition
NLP	No light perception
NMO	Neuromyelitis optica
NMOSD	Neuromyelitis optica spectrum disorders
NRS	Numeric Rating Scale
OLP	Open-label Period
ON	Optic neuritis
PCS	Physical component score
PD	Pharmacodynamic
PK	Pharmacokinetics
PLEX	Plasmapheresis; plasma exchange
PML	Progressive Multifocal Leukoencephalopathy
PT	Preferred term
RAPD	Relative afferent pupillary defect
RCP	Randomized-controlled Period
RNA	Ribonucleic acid
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Standard deviation
SFP	Safety Follow up Period
SOC	System organ class
TEAEs	Treatment-emergent adverse events
TESAEs	Treatment-emergent serious adverse events

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF APPENDICES

Amendment History

Date	Description of Change
October 18, 2018	• Amendment History: Oct 19, 2016 should be Oct 31, 2016 • Sections 2.2.1, 2.2.1.1, 2.2.1.3, 2.2.1.4 and 3.4.1.3: Introduced language related to changes in the study resulting from the Sponsor's decision to accept recommendations from the DMC, based on evidence of efficacy and safety, to stop study enrollment and allow subjects in the RCP at that time the option to enter the OLP • Section 3.1.3: Only major protocol deviation will be summarized by category. • Section 3.4.1.4: A sensitivity analysis of the primary efficacy endpoint was added to analyze the date up to and including January 27, 2017.
September 12, 2018	• Analysis of flow cytometry data was removed from Table 3.2-2 for Any MEDI-551 Population as data not being collected • Table 3.3.2.4-1: Annualized relapse rate prior to 1st IP administration instead of prior to randomization will be calculated • Table 3.3.2.4-1: Analysis of "Time (days) in between consecutive relapses, overall and by each type of relapse" was removed • Analysis of total number of T2 lesions by location was removed from Table 3.3.2.4-1 as data not being collected • Section 3.3.3: Remove "excluding placebo dose at OLP Day 15" from "Extent of Exposure" definition for more clarity • Section 3.3.4.1: Definition of concomitant medication was revised to reduce ambiguity • Section 3.4.1.4: Additional and Sensitivity Analyses of the Primary Efficacy Endpoint: Two sensitivity analyses of the primary endpoint added to 1) include SFP attack data for subjects who prematurely discontinue the RCP; and 2) censor subjects with attack onset on or before Day 15 • Section 3.4.1.6: "the negative percent agreement (NPA) =$100 \times \frac{d}{b+d}$=100% as b=0 (no AC attack determination if there is no investigator's determination)" was revised to "the negative percent agreement (NPA) =$100 \times \frac{d}{b+d}$" • Section 3.5.1.1: "percent change from baseline" was removed from the analysis of pain scale by visit as most of subjects experienced no pain at baseline. • Section 3.7.1: Typos in the Definition of Treatment-Emergent AEs for RCP period were corrected • Table 3.7.2.2-1: "TEAEs by System Organ Class (SOC)" and "TESAEs by SOC" were removed as the summary data are redundant • Table 3.7.2.2-1: Summary table "IP Related TEAEs with Severity ≥ Grade 3 by SOC and PT" was added • Section 3.7.4.4: "Change from baseline in body weight" analysis for OLP was removed as the data not being collected

	• Section 3.9: Individual MEDI-551 concentration plots by AQP4-IgG seropositive and seronegative subjects were removed
Feb 27, 2017	• Updated the study overview in Section 2.2 per Protocol Amendment 4 • Sample size was increased from 212 subjects to 252 subjects in Section 2.4 and Appendix 1 • Masked sample size reassessment was removed from Section 5 and Appendix 3 • The data cut-off date for the primary analysis was defined in Section 3.4.1.3 • The censoring date for subjects who discontinue early or who complete the RCP without an AC-determined NMO/NMOSD attack was clarified in Section 3.4.1.3 • SF-36 domain and summary score calculations were removed from Appendix 5 and stated in Section 3.5.2 that the scores will be calculated by Optum™ • Minor edits were addressed
Oct 31, 2016	• Updated the study overview in Section 3 per Protocol Amendment 3 • Clarified the masked sample size reassessment in Section 5.1 • Clarified futility stopping rules in Section 5.2. • Added sensitivity analysis to evaluate the impact of attack assessment window per FDA Type A meeting minutes in Section 3.4.1.4 • Added the time point of interest on a dosing data for vital signs data analysis in Section 3.7.4.2

1 INTRODUCTION

This document describes the statistical analyses for protocol CD-IA-MEDI-551-1155, A Double-masked, Placebo-controlled Study with Open-label Period (OLP) to Evaluate the Efficacy and Safety of MEDI-551 in Adult Subjects with Neuromyelitis Optica and Neuromyelitis Optica Spectrum Disorders.

The scope of this plan includes the primary, interim, and final analyses unless otherwise specified.

2 STUDY OVERVIEW

2.1 Study Objectives

2.1.1 Primary Objective

To compare the efficacy of MEDI-551 versus placebo in reducing the risk of a neuromyelitis optica/neuromyelitis optica spectrum disorders (NMO/NMOSD) attack in subjects with NMO/NMOSD.

2.1.2 Secondary Objectives

1. To compare the efficacy of MEDI-551 versus placebo on the reduction of Expanded Disability Status Scale (EDSS) worsening in subjects with NMO/NMOSD.
2. To compare the efficacy of MEDI-551 versus placebo on the change from baseline of low-contrast visual acuity score in subjects with NMO/NMOSD.
3. To compare the efficacy of MEDI-551 versus placebo in reducing the cumulative active MRI lesion count (new gadolinium [Gd]-enhancing or new/enlarging T2).
4. To compare the efficacy of MEDI-551 versus placebo in reducing NMO/NMOSD-related in-patient hospitalizations in subjects with NMO/NMOSD.
5. To characterize the long-term efficacy of MEDI-551 by means of annualized attack rate.
6. To evaluate the safety and tolerability of a single course of MEDI-551 in the Randomized-controlled Period (RCP) and repeated doses of MEDI-551 in the Open-label Period (OLP).
7. To characterize the pharmacokinetic (PK) profile and immunogenicity of MEDI-551 in NMO/NMOSD subjects.

2.1.3 Exploratory Objectives

1. To compare the effect of MEDI-551 versus placebo on health-related quality of life (HRQoL) as measured by the 4-week recall Short Form-36 (SF-36) Health Survey physical component score (PCS) and mental component score (MCS) in subjects with NMO/NMOSD.

2. To compare the effect of MEDI-551 versus placebo on pain as measured using the pain numeric rating scale (NRS).
3. To characterize the pharmacodynamic (PD) profile (B cells and plasma cell signature) of MEDI-551 in subjects with NMO/NMOSD.
4. To compare the effect of MEDI-551 versus placebo on aquaporin-4-antibody (AQP4-IgG) titer.
5. To compare the effect of MEDI-551 versus placebo on soluble biomarkers (eg, cytokines, chemokines, and immunoglobulins) and genomic (ribonucleic acid [RNA, microRNA]) biomarkers and other relevant cells (eg, T cells, astrocytes) in subjects with NMO/NMOSD.

2.2 Study Design

2.2.1 Overview

This is a multicenter, multinational, randomized, double-masked, placebo-controlled study with an open-label extension period to evaluate the efficacy and safety of intravenous (IV) MEDI-551 in adult subjects with AQP4-IgG seropositive and seronegative NMO and NMOSD. Following a screening period of up to 28 days, 252 subjects with NMO/NMOSD will be randomized in a 3:1 ratio (stratified by AQP4-IgG serostatus and Japan vs non-Japan origin) to receive IV MEDI-551 (300 mg at Day 1 and 300 mg at Day 15) or placebo for a period of 197 days (RCP).

Enrollment of subjects into the study will stop when a total of 67 AC-determined NMOSD attacks occur, when 252 subjects have been randomized, or following a recommendation by the independent Data Monitoring Committee (DMC) to stop enrollment, whichever occurs first.

Subjects who complete the RCP without experiencing an Adjudication Committee (AC)-determined NMO/NMOSD attack will be given the option to enroll into an OLP and will initiate or continue treatment with MEDI-551. Subjects who experience an AC-determined NMO/NMOSD attack during the RCP will be given the option to enroll into the OLP following rescue therapy. Subjects for whom the NMO/NMOSD attack is not determined by the AC will continue in the RCP until Day 197 (or until an AC-determined attack occurs). Subjects who are in the RCP when a total of 67 AC-determined NMOSD attacks have occurred or subjects who are in the RCP when enrollment is discontinued upon recommendation of the independent DMC based on evidence of efficacy and safety, will discontinue the RCP as soon as possible, preferably within 14 days and be given the option to enroll into the OLP.

The OLP will continue for a minimum of 1 year and a maximum of 3 years after the last subject enters the OLP, or until regulatory approval for MEDI-551 in the participating country, or until the Sponsor discontinues development of MEDI-551 in this indication, whichever occurs first. Subjects can choose to exit the OLP at any time at which point they enter the Safety Follow-up Period (SFP; unless consent is withdrawn).

All subjects will continue in the SFP to evaluate the long-term safety of the investigational product (IP).

2.2.1.1 Screening Period

Subjects with the diagnosis of NMO/NMOSD will be screened over a 28-day period to establish their eligibility to participate in the study based on the inclusion and exclusion criteria. AQP4-IgG serostatus will be determined by a central laboratory. For subjects that are found to be AQP4-IgG negative in the screening period, relevant data documenting their NMO disease will be assessed by an independent Eligibility Committee to assure that the data are consistent with the diagnosis of NMO. All subjects that fulfill eligibility criteria will then be randomized into the study; however, subjects undergoing screening at the time when the 252nd subject is randomized and dosed, subjects in screening at the time the 67th AC-determined attack is confirmed, and subjects undergoing screening at the time enrollment is terminated for any other reason, will not be randomized.

2.2.1.2 Randomization (Day 1)

Two hundred and fifty-two subjects will be randomized into the study in a 3:1 ratio to receive IV MEDI-551 or placebo. Randomization will occur on Day 1 (within 28 days of the start of screening) and will be stratified by AQP4-IgG serostatus determined by a central laboratory (in a ratio of approximately 80:20 seropositive and seronegative subjects, respectively), and by region (Japan vs non-Japan).

2.2.1.3 Randomized-controlled Period (Day 1 to Day 197)

Following randomization on Day 1, subjects will be treated with 300 mg MEDI-551 or placebo on Day 1 and Day 15. An oral corticosteroid course will be initiated on Day 1 (prednisone 20 mg/day or equivalent oral glucocorticoid) and will continue until Day 14. Tapering the oral corticosteroids will occur from Day 15 to Day 21. By Day 21, tapering must be completed.

During the RCP, subjects will be followed at scheduled study visits and by telephone interview. The planned duration of the RCP for each subject will be 197 days. All subjects completing the RCP without developing an AC-determined NMO/NMOSD attack will be

given the option to enter the OLP. Subjects who are in the RCP when a total of 67 AC-determined NMOSD attacks have occurred ,or subjects who are in the RCP when enrollment is discontinued upon recommendation of the independent DMC based on evidence of efficacy and safety, will discontinue the RCP as soon as possible, preferably within 14 days and be given the option to enroll into the OLP.

Subjects will be monitored for new or worsening symptoms related to NMO/NMOSD during scheduled study visits and with follow-up phone calls every 2 weeks between study visits (or if a scheduled visit is missed).

When a possible new or worsening symptom(s) related to NMO/NMOSD is identified, subjects will be required to inform the site. If an Assessment Visit is deemed necessary, this must be scheduled as soon as possible but within 72 hours of reporting of the symptoms to the site. At the Assessment Visit, subjects will initially undergo evaluations to determine if the symptoms are related to NMO/NMOSD; if related, the subjects will undergo further evaluations to determine if the symptoms meet at least ONE of the protocol-defined criteria for an NMO/NMOSD attack. In cases where a new or worsening symptom(s) does not meet at least one of the protocol-defined criteria for an NMO/NMOSD attack, the subject will continue in the RCP. The data related to the assessment of the symptoms that were determined by the investigator as not related to NMO/NMOSD will be sent to the AC for review.

Assessment of new symptoms or worsening of existing symptoms should be completed within 5 days to determine if an attack has occurred. Treatment of an attack should preferably be initiated after completion of the attack assessment and the determination that the protocol attack criteria have been met. However, the Principal Investigator can initiate rescue therapy at any time before full assessment is completed. Rescue therapy will be given as directed by the investigator. This may include IV corticosteroids, intravenous immunoglobulin, and/or plasma exchange.

Upon completion of the Assessment Visit, the complete set of data generated from the assessments will be sent to the AC regardless of whether an NMO/NMOSD attack was diagnosed according to the protocol criteria by the Principal Investigator. The adjudication process will be completed within 14 days (+ 3 days) from initiation of the Assessment Visit and the AC determination will be communicated to the Principal Investigator.

Subjects for whom the diagnosis of an NMO/NMOSD attack is not determined by the AC will be given the option to continue in the RCP until Day 197. Subjects for whom the

diagnosis of an NMO/NMOSD attack is determined by the AC will be given the option to enter the OLP.

In addition, subjects who experience an NMO/NMOSD attack that requires rescue treatment and meets the protocol-defined criteria, regardless of the outcome of the AC review, will undergo an Attack Follow-up Visit 28 days from Day 1 of the Assessment Visit. This visit may correspond with an OLP or SFP visit or may be scheduled separately.

If subjects do not wish to enter the OLP or leave the RCP at any point, they will continue to the SFP (unless consent is withdrawn).

2.2.1.4 Open-label Period

Subjects will be given the option to enter the OLP if they:

1. Complete 197 days of the RCP, or
2. Experience an AC-determined NMO/NMOSD attack during the RCP, or
3. Are in the RCP at the time when 67 AC-determined attacks have occurred, or
4. Are in the RCP at the time the Sponsor discontinues enrollment upon recommendation of the independent DMC based on evidence of efficacy and safety.

Subjects who discontinue the RCP for reasons other than an adjudicated NMO/NMOSD attack or the occurrence of 67 AC-determined attacks will not be eligible for the OLP. Reasons for subjects not entering the OLP will be captured. These subjects will be followed for safety in the SFP.

The first day of the OLP will be Day 1 (OLP Day 1). Upon entering the OLP for one of the three reasons outlined above, subjects will receive MEDI-551 as described in Figure 2.3-1. During the OLP, subjects will be followed at scheduled study visits and will continue on MEDI-551 therapy. The OLP will continue for a maximum of 3 years (after the last subject enters), until regulatory approval for MEDI-551 in each participating country or until the Sponsor discontinues development of MEDI-551 in this indication, whichever occurs first. Subjects can choose to exit the OLP at any time for any reason, including seeking alternative treatment options, at which point they enter the SFP (unless consent is withdrawn).

Subjects will be followed for NMO/NMOSD attacks in the same fashion as in the RCP and events will be centrally adjudicated.

2.2.1.5 Safety Follow-up Period

The SFP will start when a subject prematurely discontinues from the RCP or OLP. The length of the SFP will be determined by the time elapsed from the time of the last dose of the investigational product to the time of the premature discontinuation, to complete a total of 52 weeks. During the SFP, the subject will be monitored for AEs/serious adverse events (SAEs), B-cell levels, ADAs, and immunoglobulin levels.

During the SFP, a subject may receive standard treatment for their NMO/NMOSD at the discretion of the investigator.

2.3 Treatment Assignment and Masking

Randomized-controlled Period

Prior to randomization, subjects will be stratified based on AQP4-IgG serostatus (determined at screening) and region (Japan vs non-Japan). Within each stratum, subjects will be randomized in a 3:1 ratio to receive either MEDI-551 300 mg or placebo at Day 1 and 15.

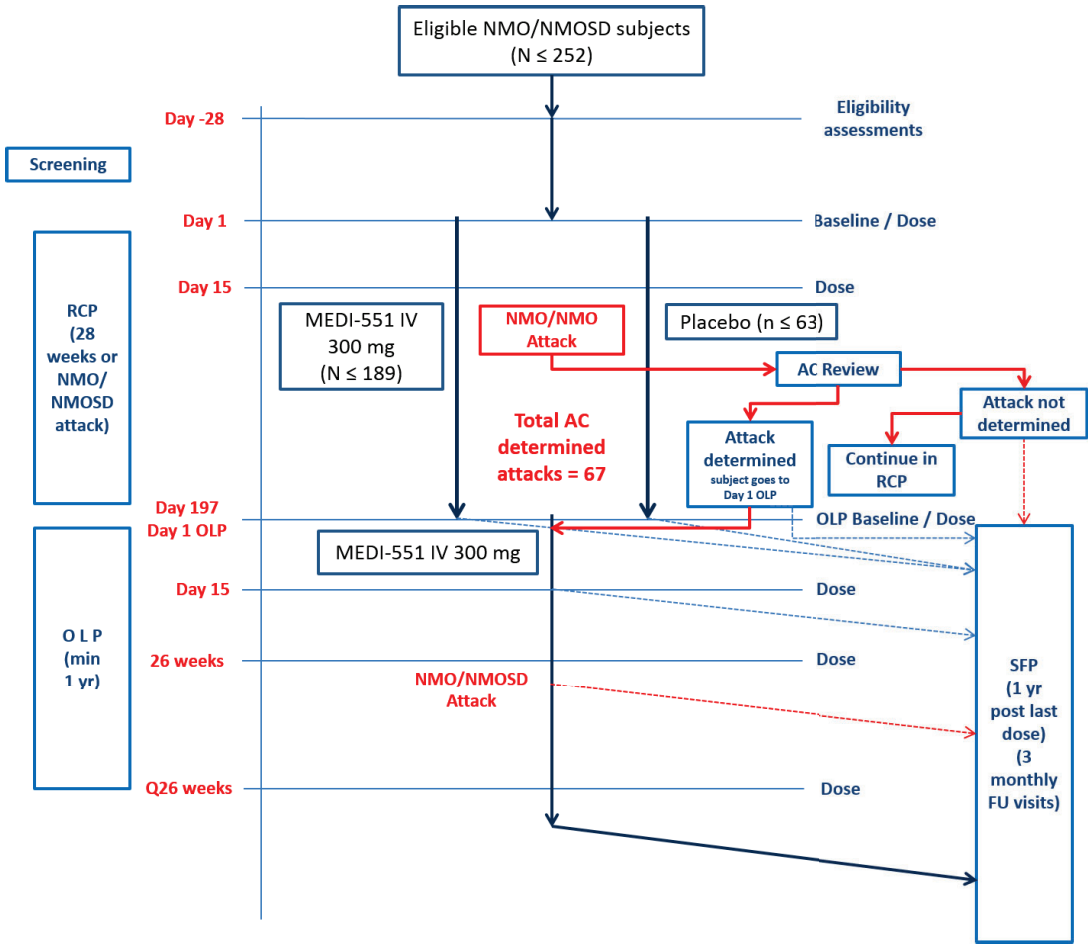

Figure 2.3-1 Study Flow Diagram

AC = Adjudication Committee; AQP4-IgG = aquaporine-4 immunoglobulin G; FU = follow-up; IV = intravenous; max = maximum; min = minimum; NMO/NMOSD = neuromyelitis optica/neuromyelitis optica spectrum disorders; OLP = Open-label Period; RCP = Randomized-controlled Period; Q26W = every 26 weeks; SFP = Safety Follow-up Period; yr = year

Open-label Period

All eligible subjects will receive open-label MEDI-551 300 mg every 26 weeks. However, subjects randomized to placebo during the RCP will receive an additional 300 mg MEDI-551 dose at Day 15 of the OLP to maintain a total initial dose of 600 mg. To protect masking, subjects randomized to MEDI-551 during the RCP will receive a placebo infusion at Day 15.

2.4 Sample Size

Study CD-IA-MEDI-551-1155 is being planned to detect a target relative reduction of 60% in risk for time from Day 1 to onset of an adjudicated NMO/NMOSD attack on or before Day 197 with 90% power and $\alpha = 0.05$ (two-sided). A total of 67 AC-determined NMO/NMOSD attacks are required for the ITT population. Subjects will be randomized in a 3:1 ratio to receive either MEDI-551 or placebo within both AQP4-IgG seropositive and seronegative strata. The stratification ratio is anticipated to be approximately 80:20 with higher allocation to the seropositive cohort. If the sero-positive cohort has 80% of the attacks, the study will have approximately 82% power to detect the target relative reduction of 60%.

Two hundred and fifty-two subjects will be randomized. This sample size is based on a blinded review of the attack rate for the first 78 subjects to complete the RCP and on a simulation based on these subjects that indicates a 90% probability of achieving the required 67 AC-determined attacks with 252 subjects.

Within each AQP4-IgG stratum, randomization will also be stratified by region (Japan vs non-Japan). The number of Japanese subjects will be determined primarily by feasibility and will not depend on a minimum number of NMO/NMOSD attacks to be observed from Japanese subjects.

3 STATISTICAL METHODS

3.1 General Considerations

All statistical calculations will be primarily performed using SAS. All data will be presented in the form of listings sorted by treatment and subject number. Tabular summaries will be presented primarily based on the following grouping:

- Stratification: ‘AQP4-IgG seropositive’, ‘AQP4-IgG seronegative’ and ‘All subjects’
  - Treatment: Placebo and MEDI-551 300 mg (nested within each level of stratification)

For pretreatment data (demographics, baseline characteristics, medical/disease history, etc.), an 'overall' group combining both treatment groups will also be added to the presentation.

Categorical data will be summarized by the frequency counts and percentage of subjects in each category. Percentages will be calculated based on non-missing observations where applicable. Continuous variables will be summarized by descriptive statistics including number of observations, mean, standard deviation (SD), median, minimum, and maximum. In general, unless stated otherwise, baseline will be defined as the last value prior to first dose of IP. Missing baseline evaluations will not be imputed and will be considered missing.

Unless otherwise stated, all efficacy analyses will be done with a two-sided test at 5% significance level. Where appropriate model based estimates will be presented with their two-sided 95% confidence interval (CI) and p-values. Categorical data analyses using the Pearson chi-square test will be substituted by the Fisher's exact test if greater than 20% of expected contingency table cell counts are less than 5.

Table 3.1-1 provides an overview of analysis milestones for this study.

Table 3.1-1 Overview of Analysis Milestones

Type of Analysis	Analysis milestones	Data to be analyzed
Unmasked futility analysis	When approximately 34 AC-determined attacks occur are observed (See Section 5)	Cumulative data on attack
Primary analysis	All subjects complete the RCP	All the data (including OLP and SFP) available at the time of database snapshot
Annual update of efficacy and safety	On an annual basis after the primary analysis	Cumulative safety and efficacy data (all study periods)
Final analysis	All subjects complete the SFP (or withdraw)	Cumulative safety and efficacy data

3.1.1 Study Points of Reference

Table 3.1.1-1 provides definition of baseline and study reference days for reporting purposes.

Table 3.1.1-1 Definition of Baseline and Study Day for Analysis Purposes

	RCP	OLP	Any MEDI-551 exposure
Baseline	Baseline is defined as the last non-missing measurement for the endpoint of interest taken before the first dose of IP. In cases where baseline measurements are taken on the same day as IP and no times are reported, it will be assumed that these measurements are taken prior to IP being administered. For subjects who are randomized but not dosed after randomization, the baseline of the study is defined as the last non-missing measurement prior to or on the date of randomization.	No separate Baseline will be defined at entry of Open label period.	For analysis by any MEDI-551 exposure, baseline for Placebo subjects may be redefined as last non-missing value prior to 1st MEDI-551 administration.
Day 1 of the study period	Day 1 for the RCP is defined as the day of first IP administration and will be denoted as Day 1 (RCP). For subjects who are randomized but not dosed after randomization, Day 1 is defined as the date of randomization.	Day 1 for the OLP is defined as the day of first open-label IP administration and will be denoted by Day 1 (OLP).	Day 1 for the MEDI-551 is defined as the day of first MEDI-551 administration irrespective of RCP or OLP and will be denoted by Day 1 (MEDI-551).
Study day for each period for reporting purpose	RCP Study Day is defined as the number of days from Day 1 (RCP) calculated as follows: = Date of Interest – Date of Day 1 (RCP) (for dates prior to Day 1 (RCP)) = Date of Interest – Date of Day 1 (RCP) + 1 (for dates on or after Day 1 (RCP))	OLP Study Day is defined as the number of days from Day 1 (OLP) calculated as follows: = Date of Interest – Date of Day 1 (OLP) (for dates prior to Day 1 (OLP)) = Date of Interest – Date of Day 1 (OLP) + 1 (for dates on or after Day 1(OLP))	Any MEDI-551 exposure is defined as the number of days from Day 1 (MEDI-551) calculated as follows: = Date of Interest – Date of Day 1 (MEDI-551) (for dates prior to Day 1 (MEDI-551)) = Date of Interest – Date of Day 1 (MEDI-551) + 1 (for dates on or after Day 1(MEDI-551))

3.1.2 Analysis Windows

Since the actual study visit for a subject may not exactly coincide with their targeted visit date, the actual visit date will be mapped to the analysis visit as provided in Table 3.1.2-1 below.

Table 3.1.2-1 Analysis Window for Randomized- controlled Period

Study Week	Nominal Day	Analysis window	Interval (days)
0	1	---	---
1	8	2-11	10
2	15	12-22	11
4	29	23-43	21
8	57	44-71	28
12	85	72-99	28
16	113	100-134	35
22	155	135-176	42
28	197	177-Day prior to OLP or SFP	---

Different sets of visit windows may be derived for measurements assessed at less frequent visits than shown in Table 3.1.2-1. Similar analysis windows will be created for OLP related analyses. If more than one actual visit falls within the same defined window, the closest visit with non-missing result to the target day will be considered for analysis. If two assessment actual dates are equidistant from the target day, the latter visit will be considered for analysis.

For retest values of lab data, the retest value (the last observation within the same visit day) will be chosen. Analysis windows will be applied after selection of such values.

3.1.3 Protocol Deviations

Protocol deviations will be recorded and will be classified as under the following categories:

- Did not fulfill eligibility criteria
- Met discontinuation criteria but continued IP treatment
- Received incorrect IP/dose
- Received prohibited concomitant medication
- Protocol-required procedure not adhered to
- Other (GCP deviations, SAE reporting deviations)

The complete specific criteria defining any protocol deviations will be agreed prior to unmasking the study. All protocol deviations will be listed and the number (%) of subjects will be summarized by treatment groups for each category for major protocol deviations.

3.2 Analysis Populations

The analysis populations are defined in Table 3.2-1. The number of subjects (without percentage) by stratification group and treatment group will be tabulated for each of these analysis populations.

Table 3.2-1 Analysis Populations

Population	Description
Intent-to-treat (ITT) population	Subjects who are randomized and receive any IP will be included in the ITT population to be analyzed according to their randomized treatment group regardless of whether subjects receive an IP different from that they were randomized to.
As-treated population	Subjects who receive any IP will be included in the as-treated population, analyzed according to treatment received. Specifically, subjects randomized to MEDI-551 who received all placebo doses will be included in the placebo group; conversely, subjects randomized to placebo who received at least one dose of MEDI-551 will be included in the active treatment group.
Open-label population	Subjects who receive at least one dose of MEDI-551 during the OLP will constitute this population.
Any MEDI-551 population	Subjects who receive at least one dose of MEDI-551 (either in the RCP or OLP).
Non-OLP population	Subset of as-treated population who do not roll over to OLP

Table 3.2-2 provides an overview of different type of analyses (demographics, safety and efficacy) by analysis populations and study periods.

Table 3.2-2 Overview of Analysis Domains by Analysis Populations and Study Periods

Period	Population				
	ITT	As-Treated	Non-OLP	Open-label	Any MEDI-551
RCP	• Demographics/ Baseline characteristics • Subject disposition • Medical and disease history • Physical/Neurological Examination • Prior and concomitant medication, steroid tapering • Efficacy, Rankin scale • HCRU, SF-36, Pain	• Exposure to investigational product • Adverse events • C-SSRS • Safety Lab/Flow cytometry lab/IgG • Immunogenicity • Vital sign, weight, ECG	---	• Demographics/Baseline characteristics • Medical and disease history • Physical/Neurological Examination • Prior Medication The information is based on baseline of RCP but will be presented only for Open-label population	
SFP	---		• Subject disposition • Concomitant medication • Adverse events • C-SSRS • Flow cytometry lab/IgG • Immunogenicity • Vital signs		
OLP	---			• Subject disposition • Concomitant medication • Efficacy, Rankin scale • HCRU, SF-36, Pain • Exposure to investigational product • Adverse events • C-SSRS • Safety Lab/Flow cytometry lab/IgG	

Table 3.2-2 Overview of Analysis Domains by Analysis Populations and Study Periods

Period	Population				
	ITT	As-Treated	Non-OLP	Open-label	Any MEDI-551
				• Vital sign	
Any MEDI-551 exposure	---				• Exposure to investigational product • Adverse events

3.3 Study Subjects

3.3.1 Subject Disposition and Completion Status

A summary of subject disposition will be presented using the categories presented in Figure 3.3.1-1:

In addition, number of subjects randomized by country and study site will be tabulated. A Kaplan-Meier (KM) curve of the time to study discontinuation will be provided. Number (%) of deaths (along with the reasons) will be reported.

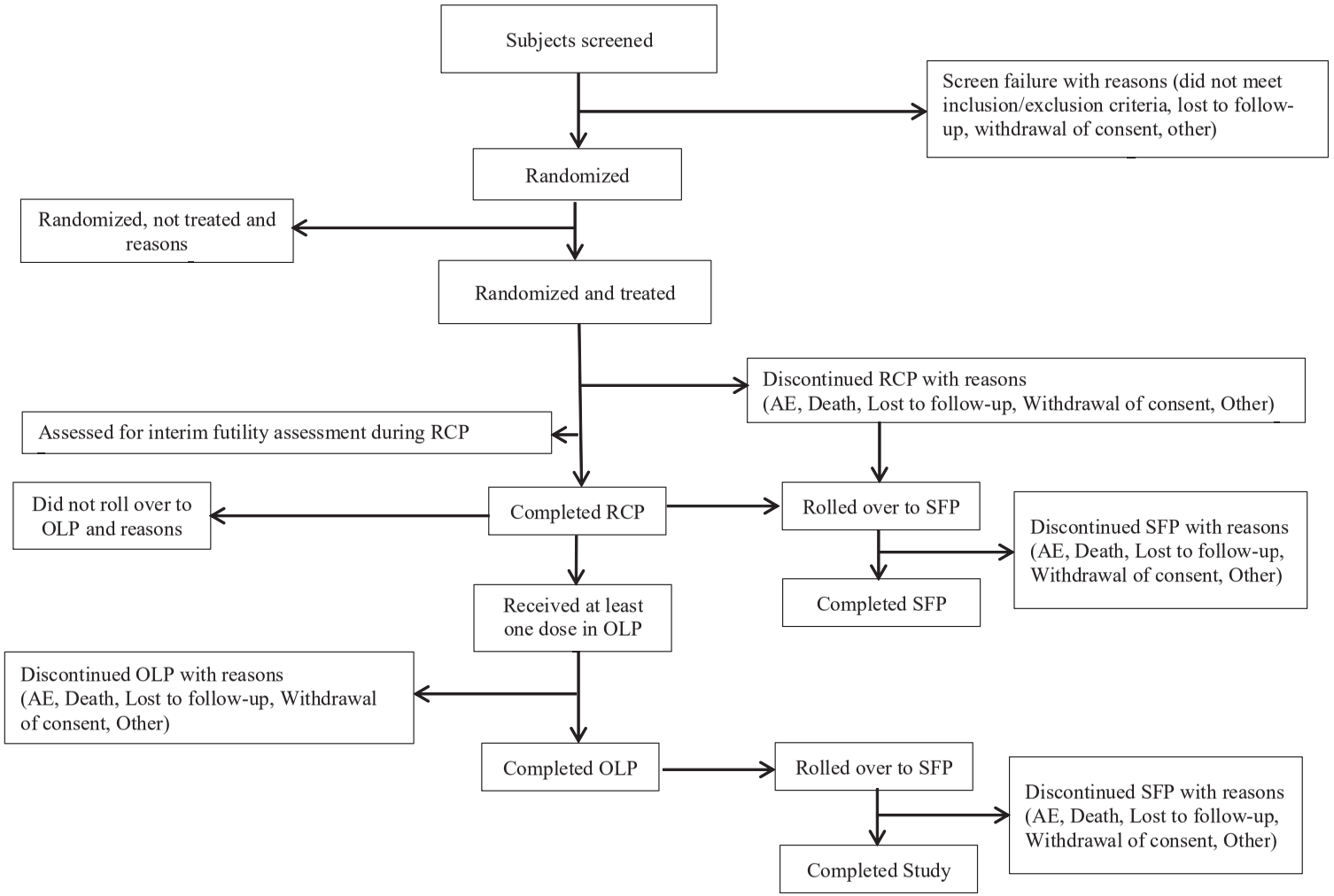

Figure 3.3.1-1 Subject Disposition

3.3.2 Demographics and Baseline Characteristics

The following demographics and baseline characteristics will be summarized for Intent-to-Treat (ITT) population by stratification factor and treatment groups as described in Section 3.1.

3.3.2.1 Demographics

- Age (years)
- Age category (< 65, ≥65 years)
- Sex (Male and Female)
- Race (American Indian or Alaskan Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Other)
- Ethnicity (Hispanic vs non-Hispanic).
- Region (US vs Non-US) and (Japan vs non-Japan)
- Height (cm)
- Weight (kg)
- BMI (kg/m^2)
- BMI category (< 18.5, 18.5-24.9, 25.0-29.9, ≥ 30)

3.3.2.2 Baseline Characteristics

Table 3.3.2.4-1 provides the list of baseline characteristics to summarize along with type of summary.

3.3.2.3 Physical Examination

Presence of abnormal physical examination finding by body system will be listed only as these are collected only at an Assessment Visit.

3.3.2.4 Neurological Examination

Presence of significantly abnormal neurological finding will be summarized by Number (%) of subjects in each category for ITT population.

Table 3.3.2.4-1 Baseline Characteristics

• Type of original diagnosis [NMO, NMOSD] - n (%) • Disease duration (years) defined as the years between date of diagnosis and date of 1st administration of IP, calculated as (date of 1st IP administration – date of diagnosis + 1)/365.25 - descriptive statistics • Disease duration category (< 5 years, ≥ 5 years) - n (%) • Type of 1st relapse [ON, Myelitis, Brain/Brainstem] - n (%) • Most recent NMO relapse type prior to 1st IP administration [ON, Myelitis, Brain/Brainstem] - n (%) • Number of prior relapse category [1, 2, 3, 4 etc.] – n (%) • Annualized relapse rate prior to 1st IP administration, calculated as total number of relapses (irrespective of type) normalized by total person-years; the person-year for each subject will be defined by the time (years) since 1st NMO/NMOSD relapse as defined below - descriptive statistics • Time (years) since 1st NMO/NMOSD relapse defined as the years between date of 1st relapse and date of 1st administration of IP, calculated as (date of 1st IP administration – date of 1st relapse + 1)/365.25 - descriptive statistics • Annualized relapse rate prior to 1st IP administration based on relapse type reported (ON, Myelitis, Brain/Brainstem) - descriptive statistics • Tested for AQP4-IgG prior to randomization (yes/no) - n (%) • AQP4-IgG seropositive prior to randomization - n (%) • Age (years) at diagnosis- descriptive statistics • Age (years) at 1st relapse- descriptive statistics • Smoking status • Never, Ever - n (%) • If ever smoked, current and former but not current - n (%) • If former but not current smoker, number of years since cessation - descriptive statistics • If current smoker, smoking duration (years) - descriptive statistics • If ever smoked, • Average number of cigars/day - descriptive statistics • Average number of pipes/day- descriptive statistics	• EDSS - n(%), descriptive statistics • EDSS category [0, 1-5, > 5] - n (%) • Functional Systems Scores - n(%), descriptive statistics • Monocular visual acuity score [Low and high contrast], left (OS) and right (OD) eye – descriptive statistics • Binocular visual acuity score [Low and high contrast] – descriptive statistics • Counting Fingers (CF), Hand Movement (HM), Light Perception (LP), No Light Perception/Total Blindness (NLP) – n (%) • RAPD (OS), RAPD (OD), No RAPD – n (%) • Pupil response score (OD and OS) (1+, 2+, 3+, 4+ and 5+) – n (%) • AQP4-IgG titer level based on flow cytometry assay (FACS) • Total number of Gd-enhancing lesions, by locations (Brain, Brainstem, Spinal cord, Optic nerve (left and right eye combined)) and overall – descriptive statistics • Co-existing autoimmune disease category [Active, Asymptomatic, Remission] (Rheumatoid arthritis, Autoimmune thyroiditis, Celiac disease, Systemic lupus, Sjögren's syndrome, Insulin dependent diabetes mellitus (IDDM), Myasthenia gravis, Pernicious anaemia, Other) – n(%) • Modified Rankin scale (mRS) score (0, 1, 2, 3, 4, 5/6) – n (%) • Worst Pain level during past 24 hours (eyes, upper back, lower back, arms, and legs) - descriptive statistics • SF-36 (4-week recall) domain scores – descriptive statistics • Columbia-Suicide Severity Rating Scale (C-SSRS) categories (see Section 3.7.4.1 for definition) for both lifetime and baseline (past 6 months) – n (%) • Chest X-ray [normal, abnormal] – n (%) • Vital signs (see section 3.7.4.2) – descriptive statistics, n (%) • Electrocardiogram (ECG) (see section 3.7.4.3) – descriptive statistics, n (%) • TB test results [Positive, Negative, Indeterminate] – n (%)

Table 3.3.2.4-1 Baseline Characteristics

• Average number of packs/day- descriptive statistics • Number of pack-years- descriptive statistics • Current Alcohol consumption status [Yes, No] - n (%) • If current consumer, number of days/week- descriptive statistics	

3.3.3 Investigational Product Exposure

Descriptive statistics of the investigational product exposure will be presented based on the measures outlined in Table 3.3.3-1.

Table 3.3.3-1 Investigational Product Exposure Summary

Population (Period)		
As-treated population (RCP only)	**Open-label population (OLP only)**	**Any MEDI-551 population (Any MEDI-551 exposure only)**
• Extent of exposure (in days) – descriptive statistics • Number of doses [1, 2]-n (%) • Dose amount (mg) – descriptive statistics • Dose amount adjusted for weight (mg/kg) based on most recent weight – descriptive statistics	• Extent of exposure (in days) – descriptive statistics • Extent of exposure (in days) categories defined by every 26 weeks or 183 day interval – n (%) • Number of MEDI-551doses excluding placebo dose at OLP Day 15 [1, 2, 3, 4, 5 etc. and mean/subject] – n(%) and descriptive statistics • Received placebo dose at Day 15 – n (%)	• Extent of exposure (in days) – descriptive statistics • Extent of exposure (in days) categories defined by every 26 weeks or 183 day interval - n (%) • Number of MEDI-551doses excluding placebo dose at OLP Day 15 [1, 2, 3, 4, 5 etc. and mean/subject] – n(%) and descriptive statistics
Extent of exposure definition		
Last RCP dose date – 1st RCP dose date + 60[a]	Last MEDI-551 dose date (OLP) – 1st OLP dose date + 60 [a]	Last MEDI-551 dose date – 1st MEDI-551dose date + 60 [a]
Person-year definition for individual subject [b]		
(Date of last day before OLP or SFP – 1st RCP dose date +1)/365.25	(Date of last day before SFP – 1st OLP dose date + 1)/365.25	(Date of last day before SFP – 1st MEDI-551 dose date + 1)/365.25

[a] based on 5 half-lives

[b] total person-years will be calculated as the sum of the person-years for individual subject

If entire dose was administered, the dose amount (mg) at a specific visit will be either 0 or 300 mg depending on what treatment the subject had received. If not, the dose will be estimated based on actual volume administered. For example, if a subject receives 175ml (instead of the full 250 mL), the dose amount (mg) will be estimated as = $300 \times \frac{175}{250} = 210$ provided the subject received MEDI-551.

3.3.4 Prior and Concomitant Medications

Number (%) of subjects who received prior and concomitant medications data will be summarized by AstraZeneca Drug Dictionary (AZDD) preferred term (PT). At each level of summarization, a subject is counted once if he/she reported one or more medications at that level.

3.3.4.1 Definition of prior and concomitant medications

- Prior medications are defined as medications with a start date occurring before the first dose date during RCP.
- Concomitant medications (RCP) are defined:
  1) Medications started or stopped on or after first dose date (RCP) and before the first dose of OLP, or
  2) Medications started or stopped on or after first dose date (RCP) and before the start of SFP if entering from RCP directly into SFP,
  3) Medications started before the first dose date (RCP) and ongoing before the start of OLP, or
  4) Medications started before the first dose date (RCP) and ongoing before the start of SFP if entering from RCP directly into SFP.
- Concomitant medications (OLP) are defined as:
  1) Medications started or stopped on or after the first dose of OLP and before the start of SFP, or
  2) Medications started before the first dose of OLP and did not stop during OLP.
- Concomitant medications (SFP) are defined as:
  1) Medications started or stopped on or after the start of SFP, or
  2) Medications started before the start of SFP and ongoing during SFP.

The following summary will be provided for the prior and concomitant medications:

- Prior NMO/NMOSD treatment categories separately for acute and maintenance treatment
  - Any immunosuppression (including azathioprine, cyclosporine, cyclophosphamide, methotrexate, mitoxantrone, mycophenolate mofetil, rituximab, corticosteroids for >21 days, or other immunosuppression)
  - Any biologic agent (eg, rituximab, eculizumab, natalizumab)
  - Plasmapheresis
  - IVIg
  - No immunosuppression (except as used to treat acute NMO attacks)
- Concomitant medications excluding rescue medication but including prophylaxis and treatment for infusion related reaction
- Rescue medications including plasmapheresis related to NMO/NMOSD attack or worsening
- Details of plasmapheresis therapy
  - Number (%) of subjects with number of episodes of plasmapheresis therapies (1, 2, 3 etc.) will be summarized
  - Mean number of sessions per plasmapheresis therapy episode (1st therapy, 2nd therapy etc.) will be summarized.

After randomization subjects will receive daily oral corticosteroids (prednisone or equivalent oral glucocorticoid) according to the schedule below during RCP:

- 20 mg on Day 1 through Day 14
- 15 mg on Day 15
- 10 mg Day 16
- 7.5 mg on Day 17
- 5 mg on Days 18 and 19
- 2.5 mg on Days 20 and 21

Descriptive summary on the following will be provided:

- Total corticosteroid (mg) received based on above tapering schedule
- Daily dose (mg/day) per subject

3.4 Efficacy Analyses

3.4.1 Primary Efficacy Endpoint Analyses

3.4.1.1 Primary Efficacy Endpoint

The primary endpoint is the time (days) from Day 1 to onset of an AC-determined NMO/NMOSD attack on or before the end of RCP (Day 197/discontinuation visit). The definition of an NMO/NMOSD attack is the presence of a new symptom(s) or worsening of an existing symptom(s) related to NMO/NMOSD that meets at least ONE of the protocol-defined criteria for an NMO/NMOSD attack provided in Table 3.4.1.4-1. This table is displayed only to pre-specify the severity grading attached to each criterion which is not presented in the protocol to eliminate any possible bias by the site when assessing for an attack. Onset day is defined as when the subject visits the site for possible attack assessment (ie, the initiation of attack assessment visit).

Per protocol requirement, sites are required to schedule an Assessment Visit with the subject within 72 hours of symptoms being reported or identified and all assessments are to be completed within 5 days of this visit. However, not all AC-determined attacks may fall within these time windows; hence a broader window will be used for primary analysis. **Only AC-determined attacks with an assessment visit scheduled within 120 hours (5 days) of reporting symptoms AND all necessary assessments done within 10 days of an Assessment Visit will be included in the primary analysis.** Other AC-determined attacks not included in primary analysis will be included in other supportive analyses.

3.4.1.2 Handling of Dropouts and Missing Data

No imputation will be performed for missing data for primary analysis.

3.4.1.3 Primary Efficacy Analysis

The following are the null and alternative hypotheses associated with the primary endpoint:

$$H_{01}: HR^{+} = 1 \text{ vs } H_{11}: HR^{+} \neq 1,$$

$$H_{02}: HR^{ALL} = 1 \text{ vs } H_{12}: HR^{ALL} \neq 1,$$

where HR^{+} indicates the hazard ratio of AC-determined NMO/NMOSD attack for MEDI-551 relative to placebo in AQP4-IgG seropositive subjects. HR^{ALL} denotes the same quantity for the ITT population. The treatment effect in the seropositive cohort will be defined as the relative reduction in HR^{+}; ie, 100 × (1-HR^{+}). A value of $HR^{+} < 1$ or a positive value of relative reduction will indicate the efficacy of MEDI-551 compared to placebo in the seropositive cohort. Similar interpretations will hold for HR^{ALL}. For the current study, the target treatment effect of 60% implies a hazard ratio (HR) of 0.4.

For the AQP4-IgG seropositive cohort, the treatment effect will be assessed using the Cox proportional hazards model with treatment indicator (MEDI-551 or placebo) as an explanatory factor; whereas for the ITT population, the model will also include serostatus as an additional explanatory factor. The HR of MEDI-551 versus placebo will be estimated together with its associated 95% CI. The SAS PROC PHREG will be used for fitting this model. Subjects who complete the Day 197 visit of the RCP or discontinue the study before Day 197 for reasons other than an AC-determined NMO/NMOSD attack will be censored in this model at the time of the Day 197/discontinuation visit.

An indicator variable for Japan versus non-Japan region will not be included in the Cox regression model due to the low number of anticipated Japanese subjects. An evaluation of consistency of treatment effect in Japanese subjects with that observed in the global study will be determined.

The data cutoff date for the primary analysis will be when the last subject completes the discontinuation visit following the 67th AC-determined attack, or after all subjects complete the RCP if 67 AC-determined attacks do not occur, or when the last subject completes the discontinuation visit following discontinuation of enrollment upon recommendation of the DMC based on evidence of efficacy and safety. Subjects who are in the RCP when a total of 67 AC-determined NMOSD attacks have occurred, or when enrollment is discontinued upon

recommendation of the DMC based on evidence of efficacy and safety, will discontinue the RCP as soon as possible, preferably within 14 days.

The procedure for the Type I error control for the primary analysis of AQP4-IgG seropositive subjects, the primary analysis of ITT population and the key secondary endpoints is described in Appendix 3.

3.4.1.4 Additional and Sensitivity Analyses of the Primary Efficacy Endpoint

An estimate of the cumulative event rates at the end of the RCP and median time to event will be calculated using the KM method. Graphical displays of KM estimates and cumulative incidence estimates will be presented by treatment group.

Following sensitivity analyses of the primary endpoint will be performed:

1. Analysis using the Cox regression model with following baseline characteristics and treatment as explanatory variables:
    - Number of prior NMO/NMOSD relapses
    - Baseline EDSS score

2. Analysis similar to primary analysis (attacks from all strata) based on unanimous adjudicated attacks (all three adjudicators agree) as events, remaining subjects will be considered as censored.

3. Analysis similar to primary analysis (attacks from all strata) including subjects who prematurely discontinue the RCP without experiencing an AC-determined attack as treatment failures (events), remaining subjects will be considered as censored.

4. Analysis similar to primary analysis (attacks from all strata) including SFP data for subjects who prematurely discontinue the RCP without experiencing an AC-determined attack. Attack data from the SFP up to 204 days after the subject's initial randomisation will be used in the analysis for subjects who prematurely discontinue the RCP. If the subject did not have an attack in the SFP before or on Day 204 then they will be censored at Day 204.

5. Analysis similar to primary analysis (attacks from all strata) based on only clinical criteria (ie excluding attacks that require MRI dependent criteria), remaining subjects will be considered as censored.

6. The impact of the loss of attacks outside of the attack assessment windows in the primary analysis will be evaluated. The AC-determined attacks falling into different attack assessment windows will be analysed similarly to primary analysis.
    a. Attacks regardless of attack assessment window;

   b. Attacks with an assessment visit scheduled within 72 hours (3 days) of reporting symptoms and all assessments done within 5 days of an Assessment Visit;
   c. Attacks with an assessment visit scheduled within 72 hours (3 days) of reporting symptoms regardless whether all assessments are done within 5 days of an Assessment Visit;

7. The impact of attacks with onset before the full pharmacodynamic effect of MEDI-551 has been reached will be evaluated. Subjects with AC-determined attacks with onset on or before Day 15 will be censored at the time of the attack.

8. Analysis similar to primary analysis including AC-determined attacks up to and including January 27, 2017, remaining subjects will be censored on January 27, 2017 (the Sponsor received a recommendation from IDMC to stop enrollment and conclude the RCP of the study as early as January 27, 2017).

Consistency of treatment effect observed in Japanese subjects with global treatment effect may be assessed by the ratio of hazard ratios (HR) as HR(Japan)/HR(Global), where respective HR estimate is obtained from Cox proportional hazards model based on respective population.

Table 3.4.1.4-1 Protocol-defined Criteria for an NMO/NMOSD Attack with criteria based severity

Attack Type	Protocol-defined Attack Criteria (with severity grading)
ON	1. Greater than 15-character drop in high-contrast Landolt C Broken Rings Chart from last visit as measured in a previously affected eye and no other ophthalmological explanation (Severe) 2. At least 2-step drop in CF to NLP from last visit as measured in a previously affected eye and no other ophthalmological explanation (Severe) 3. At least 7 or more character drop in low-contrast Landolt C Broken Rings Chart from last visit as measured in either eye alone (monocular) **AND** a new RAPD in affected eye (Moderate) 4. At least 7 or more character drop in low-contrast Landolt C Broken Rings Chart from last visit as measured in either eye alone (monocular) **AND** loss of a previously documented RAPD in fellow eye (Moderate) 5. At least 5 or more character drop in high-contrast Landolt C Broken Rings Chart from last visit as measured in either eye alone (monocular) **AND** a new RAPD in affected eye (Mild) 6. At least 5 or more character drop in high-contrast Landolt C Broken Rings Chart from last visit as measured in either eye alone (monocular) **AND** loss of a previously documented RAPD in fellow eye (Mild) 7. At least 1-step drop in CF to NLP from last visit as measured in a previously affected eye **AND** a new RAPD in affected eye (Moderate) 8. At least 1-step drop in CF to NLP from last visit as measured in a previously affected eye **AND** loss of a previously documented RAPD in affected eye (Moderate) 9. At least 7 or more character drop in low-contrast Landolt C Broken Rings Chart from last visit as measured in either eye alone (monocular) **AND** a new Gd-enhancing or new/enlarging T2 MRI lesion in the corresponding optic nerve (Moderate) 10. At least 5 or more character drop in high-contrast Landolt C Broken Rings Chart from last visit as measured in either eye alone (monocular) **AND** a new Gd-enhancing or new/enlarging T2 MRI lesion in the corresponding optic nerve (Mild) 11. At least 1-step drop in CF to NLP from last visit as measured in a previously affected eye AND a new Gd-enhancing or new/enlarging T2 MRI lesion in the corresponding optic nerve (Moderate)

Table 3.4.1.4-1 Protocol-defined Criteria for an NMO/NMOSD Attack with criteria based severity

Attack Type	Protocol-defined Attack Criteria (with severity grading)
Myelitis	12. At least 2-point worsening in 1 or more of the relevant (pyramidal, bladder/bowel, sensory) FSS compared to last visit (Severe) 13. At least 1-point worsening in EDSS score compared to last visit if previous EDSS score is 5.5 or more (Severe) 14. At least 1-point worsening in 2 or more of the relevant (pyramidal, bladder/bowel, sensory) FSS compared to last visit **AND** a new Gd-enhancing or new/enlarging T2 MRI lesion in the spinal cord (Moderate) 15. At least 0.5-point worsening in EDSS score compared to last visit if previous EDSS score is 5.5 or more **AND** a new GD-enhancing or new/enlarging T2 MRI lesion in the spinal cord – (Severe)
Brainstem	16. Isolated (not present at last visit) intractable nausea, vomiting, and/or hiccups lasting for greater than 48 hours AND a new Gd-enhancing or new/enlarging T2 MRI lesion in the brainstem (Severe) 17. At least 2-point worsening in 1 or more of the relevant (brainstem, cerebellar) FSS compared to last visit AND a new Gd-enhancing or new/enlarging T2 MRI lesion in the brainstem (Severe)
Brain	18. At least 2-point worsening in 1 or more of the relevant (cerebral, sensory, pyramidal) FSS (with a score of 3 or more at the current visit) compared to last visit AND a new Gd-enhancing or new/enlarging T2 MRI lesion in the brain consistent with the clinical presentation (Severe)

CF = counting fingers; EDSS = Expanded Disability Severity Score; FSS = Functional System Scores; Gd = gadolinium; HM = hand motion; LP = light perception; MRI = magnetic resonance imaging; NLP = no light perception; NMO/NMOSD = neuromyelitis optica/neuromyelitis optica spectrum disorders; ON= optic neuritis; RAPD = relative afferent pupillary defect.

3.4.1.5 Subgroup Analyses of Primary Endpoint

Consistency of treatment effect measured by primary endpoint in the following subgroups will be investigated for the ITT population:

- Sex (male vs female)
- Baseline EDSS (< 5 vs ≥ 5)
- Number of prior NMO/NMOSD relapses (< 2 vs ≥ 2)
- Disease duration category (< 5 years vs ≥ 5 years)
- AQP4-IgG serostatus (positive vs negative) as determined at screening

The nominal p-value and 95% CIs of treatment effect will be provided for each subgroup analysis. Forest plots will be generated to visually present the consistency of treatment effect in different subgroups with overall treatment effect.

3.4.1.6 Homogeneity of Attack Determination between Adjudication Committee and Investigators

The homogeneity of attack determination between the AC and investigator will be presented by the following quantities:

- The overall percent of agreement (OPA)
- The positive percent agreement (PPA)
- (Cohen's) Kappa Statistic and 95% CI and p-value

Let a, b, c and d represents the count of subjects as shown in the table below; then, $\mathrm{OPA} = 100 \times \frac{a+d}{a+b+c+d}$, $\mathrm{PPA} = 100 \times \frac{a}{a+c}$, the negative percent agreement (NPA) $=100 \times \frac{d}{b+d}$.

---		Investigator's determination	
		Attack	Non-attack
AC-determination	Attack	a	b
	Non-attack	c	d

Appendix 6 provides the details of range of Kappa Statistic expected under the proposed study design.

3.4.2 Secondary Efficacy Endpoint Analyses

3.4.2.1 Secondary Efficacy Endpoints

Endpoints 1, 2, 3, and 4 are key secondary endpoints to be considered for studywise Type I error control.

1. Worsening from baseline in EDSS at last visit during the RCP.

A subject will be considered to have a worsening in overall EDSS score if one of the following criteria is met:

   a. Worsening of 2 or more points in EDSS score for subjects with baseline score of 0.
   b. Worsening of 1 or more points in EDSS score for subjects with baseline score of 1 to 5.
   c. Worsening of 0.5 points or more in EDSS score for subjects with baseline score of 5.5 or more.

2. Change from baseline in low-contrast visual acuity binocular score measured by low-contrast Landolt C Broken Rings Chart, at last visit during the RCP.
3. Cumulative total active MRI lesions (new Gd-enhancing or new/enlarging T2) during the RCP.
4. Number of NMO/NMOSD-related in-patient hospitalizations. In-patient hospitalization is defined as more than an overnight stay
5. Annualized attack rate (total number of AC-determined NMO/NMOSD attacks normalized by person-years) during any exposure to MEDI-551.

3.4.2.2 Handling of Dropouts and Missing Data

Missing values of the secondary efficacy endpoints due to dropout or missing data will be handled as follows:

1. For worsening from baseline in EDSS endpoint, missing values will be considered as ‘worsening’ according to non-responder imputation (NRI) rule.
2. For low-contrast visual acuity endpoint, the last non-missing value will be considered for analysis.

3.4.2.3 Secondary Efficacy Endpoint Analyses

Secondary endpoints 1 through 4 presented in Section 3.4.2.1 for AQP4-IgG seropositive cohort will be analyzed as follows:

- Treatment effect for secondary efficacy endpoint based on EDSS worsening will be assessed using a logistic regression model with treatment and baseline EDSS as

explanatory variables. The percentage of subjects meeting the endpoint, odds ratios, p-value, and 95% CIs of the odds ratios will be presented.

- The treatment effect for the low-contrast visual acuity measured by change from baseline in low-contrast Landolt C Broken Rings Chart binocular scores will be assessed using an analysis of covariance model using treatment and baseline Landolt C Broken Rings Chart binocular score as explanatory variables.
- The treatment effect for the secondary efficacy endpoints based on the cumulative number of active MRI lesions and number of NMO/NMOSD-related in-patient hospitalizations will be tested using Negative Binomial regression with treatment as an explanatory variable.

Similar analyses will be performed for the ITT population by extending the above-mentioned models to include an indicator variable for serostatus as well.

See Section Appendix 3 for details regarding Type I error control.

Annualized attack rate (total number of AC-determined NMO/NMOSD attacks normalized by person-years) during any exposure to MEDI-551 will be summarized descriptively. See Table 3.3.3-1 for definition of person-years.

3.4.2.4 Subgroup Analyses of Secondary Efficacy Endpoints

Consistency of treatment effect based on the secondary efficacy endpoints in each of the following subgroups will be assessed based on the analysis methods described above for the ITT population.

- Sex (male vs female)
- Baseline EDSS (< 5 or ≥ 5)
- Number of prior NMO/NMOSD relapses (< 2 or ≥ 2)
- Disease duration category (< 5 years, ≥ 5 years)

The nominal p-value and 95% CIs of treatment effect will be provided for each subgroup analysis. Forest plots will be generated to visually present the consistency of treatment effect in different subgroups with overall treatment effect.

3.4.3 Healthcare Resource Utilization

Healthcare Resource Utilization (HCRU) information is collected for this study to quantify the impact of disease and treatment on the subject's medical facility visits. Information will be collected throughout the RCP and the OLP on all NMO-related HCRU visits as presented in Table 3.4.3-1.

Table 3.4.3-1 Healthcare Resource Utilization Information

Parameter	Frequency
Ambulance transport	Total number of times since the last visit
Hospitalization (intensive care)	Total number of days since the last visit
Hospitalization (general care)	Total number of days since the last visit
Emergency Room/Department visits	Total number of times since the last visit
Visits to primary healthcare physician	Total number of times since the last visit
Other healthcare visits (eg, physiotherapist)	Total number of times since the last visit
Home visit from a physician	Total number of times since the last visit
Home visit from a nurse	Total number of times since the last visit

For ambulance transport, the total will be derived for each subject by summing the total number of times since the last visit over the entire RCP and OLP separately. This total will be used for analysis. Similar totals will be derived for other parameters.

3.4.3.1 Analysis of Healthcare Resource Utilization

The total for each of the above HCRU-defined parameters will be summarized descriptively for each treatment group and separately for the RCP and OLP.

3.5 Patient Reported Outcomes

3.5.1 Pain Scale

Eleven-point numerical rating scales will be used to capture a subject's pain scores for pain experienced in the eyes, upper back, lower back, arms, and legs, respectively, where 0 = no pain and 10 = worst pain imaginable. The subject will be asked to rate the pain he/she has experienced over the last 7 days at each specified location separately.

3.5.1.1 Analysis of Pain Scale

Actual value and change from baseline of the pain scores for each body location (eyes, legs, arms, upper back, and lower back) will be summarized descriptively over all visits during RCP and OLP. The pain score averaged over all locations will be summarized.

Treatment effect for pain based on change from baseline at the last visit of the RCP will be tested using analysis of covariance models using treatment and baseline pain NRS scores as explanatory variables. Last non-missing value during RCP will be used for this analysis.

Another supportive analysis will be implemented using mixed effects model repeated measures (MMRM) approach. In this model, change from baseline will be dependent variable; baseline pain score, treatment and time will be used as independent variables. An

interaction term of treatment and time will also be included. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom and adjust standard errors. The treatment difference will be obtained with LSMEAN statement from SAS PROC MIXED for the treatment difference at a specified visit. An unstructured within subject error covariance matrix will be used for this model. Subjects' last value during RCP will be aligned with next closest scheduled visit in order to use all information available.

3.5.2 SF-36 Health Survey

The SF-36 Health Survey measures 8 dimensions of a subject's functional health and well-being and also provides 2 summary scores that characterize a subject's mental and physical health status. Two versions of this instrument (two recall periods, 4-weeks [standard] or 1-week [acute]) will be used for this study. The acute version will only be used at 'attack assessment visit' and 'attack follow up visit' during the RCP. The results from this version will be listed only. The standard version will be used at all visits and will be analyzed as described below. The SF-36 domain scores and physical and mental component summary scores will be calculated by Optum™.

3.5.2.1 Analysis of SF-36 Health Survey

Actual value, change and percent change from baseline of SF-36 scores will be summarized descriptively over all visits during RCP and OLP.

Treatment effect based on SF-36 PCS will be tested using an ANCOVA model using treatment and baseline SF-36 PCS score as explanatory variables. Similar analysis will be done for SF-36 MCS and the 8 remaining domains.

Another supportive analysis for PCS and MCS change scores during RCP will be implemented using mixed effects model repeated measures (MMRM) approach as described in section 3.5.1.1.

3.5.3 Modified Rankin Scale

The mRS ranges from 0 to 6, ranging from perfect health without symptoms to death as follows:

- 0 - No symptoms
- 1 - No significant disability. Able to carry out all usual activities, despite some symptoms
- 2 - Slight disability. Able to look after own affairs without assistance, but unable to carry out all previous activities

- 3 - Moderate disability. Requires some help, but able to walk unassisted
- 4 - Moderately severe disability. Unable to attend to own bodily needs without assistance, and unable to walk unassisted
- 5 - Severe disability. Requires constant nursing care and attention, bedridden, incontinent
- 6 – Dead

Categories 5 and 6 will be combined for analysis purposes.

Treatment effect based on mRS score during RCP will be evaluated by Wilcoxon-Mann-Whitney Odds (WMWodds) approach (Appendix 5). Only the last post-baseline mRS score during RCP for each subject will be used for this analysis. Other post-baseline scores will be presented descriptively using count and percentage in each mRS category (with 5 and 6 combined into a single category) and shift from baseline.

3.5.4 Other Efficacy Analyses

Table 3.5.4-1 provides an overview of additional efficacy analyses to be performed.

Table 3.5.4-1 Additional Efficacy Analyses

Type of analysis	Analysis Method	Analysis population
Cox proportional hazards model of investigator-determined NMO/NMOSD attack	Similar analysis as primary endpoint but based on investigator-determined NMO/NMOSD attack.	ITT
Time to either AC-determined NMO/NMOSD attacks or rescue therapy	An exploratory analysis of time to adjudicated AC-determined NMO/NMOSD attacks or rescue therapy (where an attack has not been determined by AC) during the RCP will be conducted by considering administration of rescue therapy as an event of interest. Subjects who never experience an NMO/NMOSD attack (determined by AC) nor received any rescue therapy will be censored for this analysis. In the instance where a subject never experiences an AC-determined NMO/NMOSD attack, the time to rescue therapy will be used for this analysis.	ITT
Type of NMO/NMOSD attacks	Number (%) of subjects with different types of NMO/NMOSD attacks (ON, Myelitis, Brain and Brainstem) (subjects will be included in multiple categories if applicable): • Based on AC-determined attacks • Based on investigator-determined attacks	ITT
Summary of individual criteria of NMO/NMOSD attack	Number (%) of subjects meeting each of the protocol-defined criteria for NMO/NMOSD attack will be summarized descriptively (subjects will be included in multiple categories if applicable): • Based on AC-determined attacks • Based on investigator-determined attacks	ITT
Criteria based severity of attack	Number (%) of subjects meeting each of the protocol-defined criteria of an AC-determined NMO/NMOSD attack will be summarized broken down by the severity [mild, moderate, severe] defined in Table 3.4.1.4-1	ITT
Attack type based severity	Number (%) of subjects meeting an AC-determined NMO/NMOSD attack will be summarized broken down by the severity [minor, major] defined in Appendix 4	ITT

Table 3.5.4-1 Additional Efficacy Analyses

Type of analysis	Analysis Method	Analysis population
Attack recovery	Number (%) of subjects meeting an AC-determined NMO/NMOSD attack will be summarized broken down by the recovery status [minor, major] defined in Appendix 4	ITT
Summary of symptoms of AC-determined NMO/NMOSD attacks	Number (%) of subjects broken down by different type of NMO symptoms who experienced an AC-determined attack. Subjects will be included in multiple categories if applicable. A separate summary of the percentage based on number of AC-determined attack will also be provided.	ITT
Concordance between NMO relatedness assessment between AC and investigator	Number (%) of cases by relatedness to NMO [related, not related] for each reported symptom where AC agrees or disagrees with investigators' assessment. The denominator will be the number of cases reported rather than subjects.	ITT
EDSS	Descriptive summary of change from baseline by visit	ITT and OL
EDSS category	Number (%) of subjects for EDSS category [0, 1-5, > 5] by visit	ITT and OL
EDSS worsening	Number (%) of subjects meeting EDSS worsening by visit	ITT and OL
Functional Systems Scores	Descriptive summary of change from baseline by visit	ITT and OL
MRI	Descriptive summary MRI lesion by location (Optic nerve, Spinal cord, Brain and Brainstem) and overall during RCP and OLP • cumulative new Gd-enhancing lesions • cumulative new/enlarging T2 lesions • cumulative active lesions (sum of cumulative new Gd lesions and new/enlarging T2 lesions that are not associated with a Gd lesion)	ITT and OL
Visual acuity score	Descriptive summary of change from baseline by visit for the following (low and high contrast): • monocular score, left (OS) and right (OD) • binocular score	ITT and OL
Other visual acuity parameters	Number (%) of subjects by visit for the following: • Counting Fingers (CF), Hand Movement (HM), Light Perception (LP), No Light Perception/Total Blindness (NLP) • RAPD (OS), RAPD (OD), No RAPD • Pupil response score (OD and OS) (1+, 2+, 3+, 4+ and 5+)	ITT and OL

Table 3.5.4-1 Additional Efficacy Analyses

Type of analysis	Analysis Method	Analysis population
Other analyses related to in-patient Hospitalizations (secondary endpoint)	Number (%) of subjects: • With at least one in-patient hospitalization • With at least one in-patient hospitalization or AC-determined attack • With number of in-patient hospitalizations (1, 2, >2) Descriptive summary: • Mean number of in-patient hospitalization per subject • Mean days with in-patient hospitalization per subject	ITT and OL

3.6 Pharmacodynamic Endpoint(s) and Analyses

3.6.1.1 Pharmacodynamic Endpoint(s)

The following PD endpoints will be analyzed:

- AQP4-IgG titer level and change from baseline
- B-cell absolute counts and percentage based on baseline
- Change from baseline in plasma gene signature

3.6.1.2 Analysis of Pharmacodynamic Endpoints

AQP4-IgG titer levels and corresponding change from baseline will be summarized by visit. Change from baseline will be analyzed based on the ANCOVA model with baseline value and treatment as explanatory variables.

Change from baseline by visit in plasma gene signature will be summarized descriptively for all subjects and by serostatus at screening.

B-cell (CD-19 and CD-20) absolute counts and corresponding percentage based on baseline will be summarized descriptively by visit.

3.7 Safety Analyses

Adverse events (AEs) will be coded to the corresponding system organ class (SOC) and preferred terms (PT) using current version of the Medical Dictionary for Regulatory Activities (MedDRA). The incidence, severity, and relationship to investigational product will be summarized. Specific AEs will be counted once for each subject for calculating percentages. In addition, if the same AE occurs multiple times within a particular subject, the highest severity and level of relationship observed will be reported. If any associations of interest between AEs and baseline characteristics are observed, additional stratified results may be presented. All treatment-emergent adverse events (TEAEs) will be summarized overall, as well as categorized by MedDRA SOC and PT.

3.7.1 Adverse Events

The definitions of TEAE provided in Table 3.7.1-1 will be used for analysis of adverse events based on different study periods. Figure 3.7.1-1 provides a visual representation of TEAEs during different study periods. The AE(RCP), AE(OLP) and AE(SFP) represent TEAEs during RCP, OLP and SFP, respectively.

Table 3.7.1-1 Definition of Treatment-Emergent AEs (TEAEs)

RCP	OLP	SFP	Any MEDI-551 Exposure
For subjects entering the OLP, a TEAE will be defined as any AE with an onset time on or after the first dose of IP in the RCP up to the time prior to the first OLP dose.	A TEAE will be defined as any AE with an onset time on or after 1st dose of OLP up to and including the day prior to the SFP entry.	A TEAE will be defined as any AE with an onset on or after entry to SFP.	A TEAE will be defined as any AE with onset on or after the first MED-551 dose (either RCP or OLP) up to and including the day prior to entry to SFP.
For subjects entering the SFP after RCP, a TEAE will be defined as any AE with an onset on or after the first dose of IP in the RCP up to and including the day prior to the entry to SFP.	---		
For subjects not entering the OLP or SFP after RCP, a TEAE will be defined as any AE with an onset on or after the first dose of IP in the RCP			

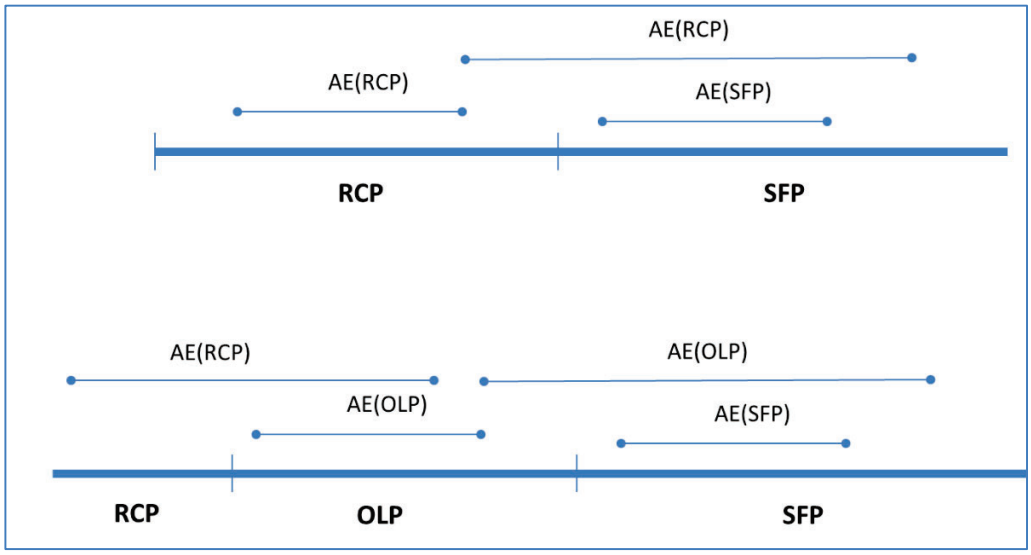

Figure 3.7.1-1 Visual definition of TEAEs during different study periods

3.7.1.1 Adverse Events of Special Interest

The following Adverse Events of Special Interest (AESI) will be summarized; the infusion related reaction, anaphylactic reaction and hypersensitivity will be collected from the case report form. Infections will be identified by the corresponding System Organ Class (SOC); for remaining AESIs, relevant search criteria will be established prior to database lock.

- Infusion Related Reaction
- Anaphylactic Reaction
- Hypersensitivity
- Infections
- Hepatic Function Abnormality
- Cytopenia
- Opportunistic Infection
- Progressive Multifocal Leukoencephalopathy (PML)

Symptoms corresponding to infusion-related reaction will be summarized separately and will not be combined with TEAE analysis.

3.7.1.2 Deaths and Treatment Discontinuations due to Adverse Events

Deaths, fatal AEs, and AEs leading to discontinuation of IP will be summarized.

3.7.2 Analysis of Adverse Events and Serious Adverse Events

Treatment-Emergent adverse events and serious adverse events during RCP, OLP, SFP and any MEDI-551 exposure will be summarized as described in Table 3.7.2.2-1.

TEAEs per 100 Person-years for a specific reporting period will be calculated as

$$= 100 \times \frac{\text{Total number of TEAEs for specific reporting period}}{\text{Total Person years for specific reporting period}}$$

See Table 3.3.3-1 for definition of person-years.

3.7.2.1 Analysis of non-NMO related TEAE

Number (%) of cases will be summarized by preferred term for RCP to establish the degree of agreement between the Adjudication Committee and the investigator's decision about

symptoms not being related to NMO and thus reported as TEAEs. The denominator will be decided by the total number of each such symptom reported.

3.7.2.2 Subgroup Analyses of Adverse Events

Treatment-Emergent adverse events (by SOC and PT) during RCP will be summarized by sex (male vs female).

Table 3.7.2.2-1 Type of TEAE Analysis by Period

Type of TEAE Analysis	RCP	OLP	SFP	Any MEDI-551 Exposure
Overall Summary of TEAEs	X	X	X	X
TEAEs by SOC and PT	X	X	X	X
IP Related TEAEs by SOC, PT	X	X		X
TEAEs with severity ≥ Grade 3 by SOC and PT	X	X		X
TESAEs by SOC and PT	X	X	X	X
TEASEs by Seriousness Criteria	X	X	X	X
TEAEs resulting in permanent discontinuation of the IP by SOC and PT	X	X		X
TEAEs resulting in death by SOC and PT	X	X	X	X
TEAEs (≥5% in MEDI-551)	X	X		X
TEAEs by SOC, PT and by Highest Severity	X	X	X	X
IP Related TEAEs by SOC, PT and by Highest Severity	X	X		X
IP Related TEAEs with Severity ≥ Grade 3 by SOC and PT	X	X		X
IP Related TESAEs by SOC and PT	X	X		X
TEAEs with Severity ≥ Grade 3 and/or TESAEs by SOC and PT	X	X		X
IP Related TEAEs with Severity ≥ Grade 3 and/or TESAEs by SOC and PT	X	X		X
TEAEs by PT sorted by frequency in MEDI-551 group	X	X		X
TEASEs by Preferred Term Sorted by Frequency	X	X	X	X
All deaths	X	X	X	X
AESI by SOC and PT	X	X	X	X
TEAEs per 100 Person-years	X	X		X
TESAEs per 100 Person-years	X	X		X
'X' represents that an analysis will be performed for this period				

3.7.3 Clinical Laboratory Evaluation

Laboratory parameters will be assessed at baseline as well as throughout the study. Laboratory toxicities will be defined based on the National Cancer Institute's Common Terminology Criteria for Adverse Events Version 4.0 (NCI CTCAE V 4.0).

Following analyses will be presented for RCP and OLP:

- Grade 3-4 Clinical Laboratory Toxicities
- Baseline Toxicity Grade in different Laboratory Categories
- Worst Toxicity Grade (during post-baseline period) in different Laboratory Categories
- Toxicity Grade by Visit in different Laboratory Categories
- At Least 2-Grade Shift from Baseline to Worst Toxicity Grade in different Laboratory Categories
- Shifts from Baseline Relative to the Normal Range in different Laboratory Categories
- Absolute values, change and percent changes in Laboratory values

3.7.4 Other Safety Evaluations

3.7.4.1 Columbia Suicide Severity Rating Scale

The C-SSRS involves a series of probing questions that inquire about possible suicidal thinking and behavior, and classifies these events of interest into 11 categories of interest as part of the assessment process. The terminology considered important includes 5 levels of suicidal ideation, 5 levels of suicidal behavior, and the category "self-injurious behavior, no suicidal intent." Different versions of the C-SSRS are administered at different visits during the study. The "Baseline/Screening" questionnaire will be administered to each subject at screening. This version of the scale combines the "Baseline" and "Screening" versions to assess suicidality in a subject's lifetime and during a predefined time period. The "Since Last Visit" version of the C-SSRS will be administered at all visits for all subjects randomized into the study.

The following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definitions of the composite and comparative endpoints, and to enable clarity in the presentation of the results.

Table 3.7.4.1-1 C-SSRS categories for reporting

Category #	Category
1	Wish to be Dead
2	Non-specific Active Suicidal Thoughts
3	Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
4	Active Suicidal Ideation with Some Intent to Act, without Specific Plan
5	Active Suicidal Ideation with Specific Plan and Intent
6	Preparatory Acts or Behavior
7	Aborted Attempt
8	Interrupted Attempt
9	Actual Attempt (non-fatal)
10	Completed Suicide
11	Self-injurious behavior without suicidal intent (not suicide-related)

The following outcome is a numerical score derived from the C-SSRS categories. The score is created at each assessment for each subject and is used for determining treatment emergence.

Suicidal Ideation Score: The maximum suicidal ideation category (1-5 on the C-SSRS) present at the assessment. Assign a score of 0 if no ideation is present.

Composite endpoints to be evaluated based on the above categories are defined in Table 3.7.4.1-2. Some of these endpoints will depend on the definition of Suicidal Ideation Score defined below.

Table 3.7.4.1-2 Composite Endpoints based on C-SSRS

Endpoint	Description
Suicidal ideation	A “yes” answer at any time during treatment to any 1 of the 5 suicidal ideation questions (Categories 1-5) on the C-SSRS.
Suicidal behaviour	A “yes” answer at any time during treatment to any 1 of the 5 suicidal behaviour questions (Categories 6-10) on the C-SSRS.
Suicidal ideation or behaviour	A “yes” answer at any time during treatment to any 1 of the 10 suicidal ideation and behaviour questions (Categories 1-10) on the C-SSRS.

Comparative endpoints to be evaluated are defined in Table 3.7.4.1-3.

Table 3.7.4.1-3 Comparative Endpoints based on C-SSRS

Endpoint	Description
Treatment-emergent suicidal ideation compared to Baseline	An increase in the maximum suicidal ideation score during treatment from the maximum suicidal ideation category at baseline (excluding "lifetime" scores from the Baseline/Screening C-SSRS scale).
Treatment-emergent serious suicidal ideation compared to Baseline	An increase in the maximum suicidal ideation score to 4 or 5 on the C-SSRS during treatment from a baseline score of 0-3 (excluding "lifetime" scores from the Baseline/Screening C-SSRS scale).
Improvement in suicidal ideation	Defined as a decrease in suicidal ideation score at the time point of interest (eg, during RCP and OLP) from baseline.
Emergence of suicidal behaviour compared to all prior history	The occurrence of suicidal behaviour (Categories 6-10) during treatment from not having suicidal behaviour (Categories 6-10) at baseline (includes "lifetime" and/or "screening" scores from the Baseline/Screening C-SSRS scale).

Analysis of C-SSRS will be restricted to subjects who have at least one response at both baseline and post-baseline visits. Number (%) of subjects meeting each of the categories defined in Table 3.7.4.1-2 and Table 3.7.4.1-3 and suicidal ideation score will be presented descriptively; no p-value will be produced. Missing data will not be imputed.

3.7.4.2 Vital Signs

Change from Baseline in the following vital sign parameters will be collected in this study: systolic blood pressure, diastolic blood pressure, body temperature, pulse rate, and respiratory rate.

Vital signs data will be summarized by visit and by treatment group using descriptive statistics. For vital sign assessments collected on the dosing day, the vital sign data will be presented by time points of interest as well. The time points of interest are as follows:

- Pre-dose assessment;
- During dose assessment: 15min, 30min;
- After dose assessment: immediately after dosing, the last assessment before discharge.

3.7.4.3 Electrocardiogram

Descriptive statistics for the following electrocardiogram (ECG) variables will be presented:

- Heart Rate (bpm), RR interval (msec), PR Interval (msec), Atrial rate (bpm), QRS duration (msec), Ventricular rate (bpm), QT interval (msec)

Corrected QT interval based on Fridericia approach, $QTc = QT/RR^{0.33}$ with QT in msec and RR in seconds will also be derived and summarized descriptively. Furthermore, following summaries will be provided:

Number (%) of subjects meeting the following criteria:

- QTc interval > 450
- QTc interval > 480
- QTc interval > 500
- QTc interval increases from baseline > 30
- QTc interval increases from baseline > 60
- Result [Normal, Abnormal, not clinically significant, Abnormal, clinically significant], as collected on case report form.

3.7.4.4 Body Weight

Change from baseline in body weight will be summarized descriptively for RCP. Shift analysis of following BMI categories will be performed:

- Underweight (<18.5 kg/m^2)
- Normal weight (18.5 kg/m^2 – 24.9 kg/m^2)
- Overweight (25 kg/m^2 – 29.9 kg/m^2)
- Obese (≥ 30 kg/m^2)

3.8 Immunogenicity

Number (%) of subjects meeting the following will be summarized related to Anti-Drug Antibody (ADA):

- ADA positive at baseline
- ADA positive any time during post-baseline
- ADA positive at both baseline and any time post-baseline
- ADA positive at any time post-baseline but not detected (or missing) at baseline
- ADA not detected at any time post-baseline and detected at baseline
- Persistent positive (defined as positive at ≥ 2 post-baseline assessments (with ≥ 16 weeks between first and last positive) or positive at last post-baseline assessment)

- Transient positive (defined as negative at last post-baseline assessment and positive at only one post-baseline assessment or at ≥ 2 post-baseline assessments (with < 16 weeks between first and last positive)
- ADA positive during post-baseline by visit

Following additional analyses will be performed to assess the impact of ADA on efficacy:

Number (%) of subjects with an adjudicated NMO/NMOSD attack during RCP

- based on baseline ADA positive subjects
- by ADA status (positive or negative) where the status is considered in a cumulative manner at each time point (i.e., if a subject had a positive sample at any prior time before an efficacy assessment visit then that subject would be counted as positive through that time point)

Number (%) of subjects TEAEs (by SOC and PT) during RCP by ADA status will be summarized.

3.9 Pharmacokinetics

MEDI-551 concentration will be summarized descriptively by visits separated by AQP4-IgG seropositive and seronegative NMO/NMOSD subjects. Mean MEDI-551 concentration versus time data will be plotted by AQP4-IgG seropositive and seronegative subjects. PK parameters will be estimated using noncompartmental methods. PK parameters will be listed by AQP4-IgG seropositive and seronegative NMO/NMOSD subjects and summarized descriptively.

Population PK or PKPD analysis may be conducted to include the exploratory analysis to identify the covariate that affect MEDI-551 PK or to investigate the relationship between MEDI-551 PK, PD and response as appropriate.

4 MULTIPLICITY ADJUSTMENT

The study is designed to strongly control the overall 2-sided Type I error rate of $\alpha = 0.05$ based on the Bonferroni-based chain procedure (Bretz et al, 2009; Millen and Dmitrienko, 2011). The primary null hypothesis will be hierarchically tested first at $\alpha = 0.05$ in the AQP4-IgG seropositive cohort, and, if significant, it will be further tested in the ITT population at $\alpha = 0.05$. If and only if the treatment group comparison is statistically significant within the ITT population, secondary hypotheses will be tested. Null hypotheses for the 4 key secondary endpoints will follow the same sequential testing strategy as the primary analysis (testing within seropositive subjects first, followed by the ITT population if the comparison

within seropositive subjects is statistically significant). Each secondary hypothesis will be initially tested based on the Bonferroni method at $\alpha = 0.05/4 = 0.0125$. If the null hypothesis for a particular secondary endpoint is rejected across both the seropositive and the ITT populations, the Type I error saved will be propagated equally to other non-rejected sets of secondary null hypotheses. The testing procedure will be repeated until all null hypotheses are rejected or no further null hypothesis can be rejected. Details are described in Appendix 3.

5 INTERIM FUTILITY ANALYSIS

An unmasked interim analysis will be conducted for futility assessment by an independent Data Monitoring Committee when approximately 34 (50% of the total planned AC-determined) NMO/NMOSD attacks occur in this study. The assessment of futility will be based on predictive power, which is based on the average conditional power calculated using the observed treatment effect trend and associated variability at the time of interim analysis to predict the final study outcome. All available subjects at the time of interim data cut will be included in the futility analysis. Details are provided in Appendix 2.

To be consistent with the hypothesis testing of the primary efficacy analysis as described in Section 3.4.1.3, the primary efficacy endpoint, time (days) from Day 1 to onset of an AC-determined NMO/NMOSD attack on or before the end of RCP, will be analyzed in the AQP4-IgG seropositive cohort as well as the ITT population for the interim futility analysis. The events to be included in the futility analysis will follow the methods described in the Section 3.4.1.1.

For the AQP4-IgG seropositive cohort, the treatment effect will be assessed using a Cox proportional hazards model with treatment indicator (MEDI-551 or placebo) as an explanatory factor; whereas for the ITT population, the model will also include serostatus as an additional explanatory factor. Subjects who have completed the RCP without having an AC-determined NMO/NMOSD attack will be censored on the date of completing RCP (Day 197 visit date). Subjects who have discontinued from the RCP for reasons other than an AC-determined NMO/NMOSD attack will be censored on the day of discontinuation from RCP (discontinuation visit date). Subjects who are ongoing in the RCP will be censored on the date of interim data cut. The SAS PROC PHREG will be used to fit this Cox proportional hazards model.

The predictive power based on the empirical trend observed at the interim analysis will be calculated by

$$\Phi\left[\frac{Z(t)-\sqrt{t}Z_{\alpha/2}}{\sqrt{1-t}}\right],$$

where $Z(t)$ is the observed Z-statistic at the interim futility analysis, t is observed information fraction and $Z_{\alpha/2}$ is upper $\alpha/2\,th$ quantile from standard normal distribution, and α=0.05 (Lan et al., 2009). The Z(t) will be calculated by the square root of the chi-squared statistic obtained from the aforementioned Cox proportional hazard model for the AQP4-IgG seropositive cohort and the ITT population, respectively. The observed information fraction *t* will be calculated using the observed number of AC-determined attacks at the interim analysis (ITT or seropositive) divided by the planned total number of AC-determined attacks for RCP (eg, 67 AC-determined attacks in the ITT population and 57 AC-determined attacks in the AQP4-IgG seropositive cohort for RCP).

The study may be declared as "futile" if

- the calculated predictive power is < 20% in the ITT population; AND
- the calculated predictive power is < 20% in the AQP4-IgG seropositive cohort.

If the study is declared futile the primary endpoint analysis and key secondary endpoint analyses will be conducted and all other endpoints will be summarized.

6 INTER AND INTRARATER RELIABILITY

A study will be conducted to establish inter- and intrarater reliability of the AC members. A random sample of 50% of AC determinations (separately for attack and non-attack) during the RCP will be selected for this purpose. For individual subject thus selected individual rater decision for both original assessment and reassessment time point will be included for this analysis. The intraclass correlation coefficient (ICC) for both inter- and intrarater reliability will be assessed by a random effects model (Vangeneugden et al.) based on the binary response (attack vs non-attack). The source of individual random effects will be attributed to subject, rater, and time (original vs reassessment). A joint random effect of rater and time will also be included to account for any excess variability. The corresponding statistical model is described below:

$$Y_{prd} = \frac{exp(\alpha + b_p + b_r + b_d + b_{rd})}{1 + exp(\alpha + b_p + b_r + b_d + b_{rd})} + \varepsilon_{prd}$$

where b_p, b_r, b_d and b_{rd} are the independent random effects due to subject, rater, time and joint random effect of rater and time distributed as normal with zero mean and variances $\sigma_p^2, \sigma_r^2, \sigma_d^2$ and σ_{rd}^2 respectively. Also ε_{prd} represents an error term normally distributed (independent of other random effects) with mean zero and variance σ_e^2. The intraclass correlation coefficients are defined as follows:

$$\mathrm{ICC_{Intra}} = \frac{\sigma_p^2+\sigma_r^2}{\sigma_p^2+\sigma_r^2+\sigma_d^2+\sigma_{rd}^2+\sigma_e^2} \text{ and } \mathrm{ICC_{Inter}} = \frac{\sigma_p^2+\sigma_d^2}{\sigma_p^2+\sigma_r^2+\sigma_d^2+\sigma_{rd}^2+\sigma_e^2}$$

The CIs of the ICCs will be calculated based on the method outlined in Carrascoa et al. The standard error of the ICC is obtained by using the delta method:

$$\mathrm{V(ICC)} = \mathbf{d}'\Sigma\mathbf{d}$$

where $\mathbf{d}$ represents the derivatives of ICC with respect to their components denoted by the vector $\boldsymbol{\theta} = (\sigma_p^2, \sigma_r^2, \sigma_d^2, \sigma_{rd}^2, \sigma_e^2)$, $\mathbf{d} = \frac{d}{d\boldsymbol{\theta}}(\mathrm{ICC})$ and Σ is the variance–covariance matrix of the $\boldsymbol{\theta}$. Let, $V_{tot} = \sigma_p^2 + \sigma_r^2 + \sigma_d^2 + \sigma_{rd}^2 + \sigma_e^2$, then $\mathbf{d}$ will be defined by the following quantities:

$$\frac{d}{d\sigma_p^2}(\mathrm{ICC_{Intra}}) = \frac{d}{d\sigma_r^2}(\mathrm{ICC_{Intra}}) = \frac{\sigma_d^2+\sigma_{rd}^2+\sigma_e^2}{V_{tot}^2} = \frac{1-\mathrm{ICC_{Intra}}}{V_{tot}},$$

$$\frac{d}{d\sigma_p^2}(\mathrm{ICC_{Inter}}) = \frac{d}{d\sigma_d^2}(\mathrm{ICC_{Inter}}) = \frac{\sigma_r^2+\sigma_{rd}^2+\sigma_e^2}{V_{tot}^2} = \frac{1-\mathrm{ICC_{Inter}}}{V_{tot}},$$

$$\frac{d}{d\sigma_d^2}(\mathrm{ICC_{Intra}}) = \frac{d}{d\sigma_{rd}^2}(\mathrm{ICC_{Intra}}) = \frac{d}{d\sigma_e^2}(\mathrm{ICC_{Intra}}) = -\frac{\sigma_p^2+\sigma_r^2}{V_{tot}^2} = -\frac{\mathrm{ICC_{Intra}}}{V_{tot}},$$

$$\frac{d}{d\sigma_r^2}(\mathrm{ICC_{Inter}}) = \frac{d}{d\sigma_{rd}^2}(\mathrm{ICC_{Inter}}) = \frac{d}{d\sigma_e^2}(\mathrm{ICC_{Inter}}) = -\frac{\sigma_p^2+\sigma_d^2}{V_{tot}^2} = -\frac{\mathrm{ICC_{Inter}}}{V_{tot}}.$$

The normal approximation of the estimators of the ICC can be improved by using the inverse hyperbolic tangent transformation $Z = \frac{1}{2}\ln\frac{1+ICC}{1-ICC}$. Since Z is a function of an asymptotically normal statistic, the delta method can be used to obtain its asymptotic distribution. Thus $\hat{Z} \sim N\left(0, V(\hat{Z})\right)$ where $V(\hat{Z})$ is approximated as $V\left(\hat{Z}\right) = \frac{V(\widehat{ICC})}{\left(1-\widehat{ICC}^2\right)^2}$. Hence the confidence limits for ICC can be estimated by replacing the parameter with their sample counterparts to derive boundaries for Z and transform these bounds as $\frac{\exp(2Z)-1}{\exp(2Z)+1}$.

7 REFERENCES

Bedi GS, Brown AD, Delgado SR, Usmani N, Lam BL, Sheremata WA. Impact of rituximab on relapse rate and disability in neuromyelitis optica. Mult Scler. 2011 Oct;17(10):1225-30.

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. Stat Med. 2009 Feb 15;28(4):586-604

Carrascoa, Josep L; Phillipsb, Brenda R; Puig-Martineza, Josep; Kingb, Tonya S; Chinchilli, Vernon M. Estimation of the concordance correlation coefficient for repeated measures using SAS and R. Computer Methods and Programs in Biomedicine, Volume 109, Issue 3, March 2013, Pages 293–304

Costanzi C, Matiello M, Lucchinetti CF, Weinshenker BG, Pittock SJ, Mandrekar J, et al. Azathioprine: tolerability, efficacy, and predictors of benefit in neuromyelitis optica. Neurology. 2011 Aug16;77(7):659-66

Churilov L, Arnup S, Johns H, Leung T, Roberts S, Campbell BC, Davis SM, Donnan GA. An improved method for simple, assumption-free ordinal analysis of the modified Rankin Scale using generalized odds ratios Int J Stroke. 2014 Sep 4

Divine G, Norton HJ, Hunt R, Dienemann J. A review of analysis and sample size calculation considerations for Wilcoxon tests. Anesth Analg. 2013;117:699–710

Hade E, Jarjoura D, Wei L. Sample size re-estimation in a breast cancer trial. Clin Trials. 2010 7: 219-226.

Jacob A, Weinshenker BG, Violich I, McLinsky N, Kurpp L, Fox RJ, et al. Treatment of neuromyelitis optica with rituximab: retrospective analysis of 25 patients. Arch Neurol. 2008;65(11):1443-48.

Kim SH, Kim W, Li XF, Jung IJ, Kim HJ. Repeated treatment with rituximab based on the assessment of peripheral circulating memory B cells in patients with relapsing neuromyelitis optica over 2 years. Arch Neurol. 2011 Nov;68(11):1412-20.

Lan, Gordon K. K., Hu, Peter and Proschan, Michael A. A Conditional Power Approach to the Evaluation of Predictive Power. Statistics in Biopharmaceutical Research May 2009, Vol. 1, No. 2

Landis, J.R.; Koch, G.G. (1977). The measurement of observer agreement for categorical data Biometrics 33 (1): 159–174.

Millen B, Dmitrienko A. Chain procedures: a class of flexible closed testing procedures with clinical trial applications. Stat Biopharm Res. 2011; 3:14-30

Pittock SJ, Lennon VA, McKeon A, Mandrekar J, Weinshenker BG, Lucchinetti CF, O'Toole O, Wingerchuk DM. Eculizumab in AQP4-IgG-positive relapsing neuromyelitis optica spectrum disorders: an open-label pilot study. Lancet Neurol. 2013 Jun;12(6):554-62.

Vangeneugden, Tony; Molenberghs, Geert; Laenen, Annouschka; Alonso, Ariel; Geys, Helena. Generalizability in nongaussian longitudinal clinical trial data based on generalized linear mixed models. J. Biopharm. Statist. 18 (2008), no. 4, 691–712.

Appendix 1 Sample Size Calculation

The proposed study is being planned to detect a target relative reduction of 60% in risk for time (days) from Day 1 to onset of an AC-determined NMO/NMOSD attack on or before Day 197 with 90% power and $\alpha = 0.05$ (two-sided), assuming an unequal randomization ratio of 3:1 to MEDI-551 versus placebo. Using the following formula, a total of 67 AC-determined NMO/NMOSD attacks will be required:

$$E = \left[\frac{z_{1-\alpha/2} + z_{1-\beta}}{ln(HR)}\right]^2 \times \frac{(r+1)^2}{r},$$

where $z_{1-\alpha/2}$ and $z_{1-\beta}$ are $100(1 - \alpha/2)^{th}$ and $100(1 - \beta)^{th}$ percentile of a standard normal distribution, $\boldsymbol{r} = 3$ is the randomization ratio and $\boldsymbol{HR}$ = hazard ratio = 0.4.

Following the completion of the first 78 subjects in this study the sample size was reviewed. The attack rate of the first 78 subjects to complete the RCP was used to simulate the remaining subjects to estimate the probability of reaching the required 67 AC-determined attacks. The simulation showed that with a sample size of 252 subjects there is a 90% probability of reaching the required 67 AC-determined attacks.

The remaining text in this appendix describes the process of defining the original sample size of 212 subjects.

Assuming hazard rates of 1.5/year and 1.0/year for an NMO/NMOSD attack in the placebo arm for seropositive and seronegative groups, respectively, and using the following formula:

$$N = \frac{E}{\Pr(Fail)},$$

a total of 212 (with an upward adjustment) subjects are expected to be enrolled in this study where

$$\Pr(\text{Fail}) = s_1 P_1 + s_2 P_2$$

is the weighted average of the failure probabilities P_1 and P_2 in seropositive and seronegative cohorts, respectively, with associated stratum specific weights $s_1 = 0.8$ and $s_2 = 0.2$. Recall that subjects will be enrolled to seropositive and seronegative cohorts in 80:20 ratio. Furthermore, the probability of an attack within i^{th} cohort (i = 1, 2) is given by

$$P_i = (1 - w)p_{i0} + wp_{i1}$$

the weighted average of the failure probabilities p_{i0} and p_{i1} for placebo and MEDI-551 groups, respectively; $p_{i0} = 1 - e^{-\lambda_{i0}F}$, $p_{i1} = 1 - e^{-\lambda_{i1}F}$; F = 197/365 = 0.54 to indicate the 197 days (28 weeks) duration of the controlled portion of the study; note that based on the

assumptions for this study, $\lambda_{10} = 1.5, \lambda_{11} = 0.6, \lambda_{20} = 1$ and $\lambda_{21} = 0.4$ and $w = \frac{r}{1+r}$.= 0.75. Since the primary endpoint will be tested using the Cox proportional model, a simulation approach was used to investigate the power under alternative hypotheses mentioned above. Based on 5,000 simulations, the power for the overall and seropositive populations were observed as 94% and 91%, respectively. Moreover, the total required 67 events (212 subjects) were, on an average, distributed in seropositive and seronegative cohorts as 57 (168 subjects) and 10 (44 subjects) respectively.

The underlying hazard rate of time to an NMO/NMOSD attack in the seropositive placebo group is estimated based on a meta-analytic approach from several open-label cohort studies (see Table 7-1) in the absence of any placebo-controlled study in this population to date. These studies consisted of both seropositive and seronegative subjects and hence any finding related to hazard rate derived from these studies can be considered as an underestimation when applied for seropositive subgroup in the current study. Following are the 5 open-label studies considered with a total of 161 subjects:

Table 7-1 Open-label Cohort Studies in Neuromyelitis Optica

Author	Title
Bedi et al	Impact of rituximab on relapse rate and disability in neuromyelitis optica. Mult Scler. 2011 Oct;17(10):1225-30
Costanzi et al	Costanzi C, Matiello M, Lucchinetti CF, Weinshenker BG, Pittock SJ, Mandrekar J, et al. Azathioprine: tolerability, efficacy, and predictors of benefit in neuromyelitis optica. Neurology. August 16, 2011;77(7):659-66.
Jacob et al	Treatment of neuromyelitis optica with rituximab: retrospective analysis of 25 patients. Muscle Nerve. 2008 Jan;39(1):87-90
Kim et al	Kim SH, Kim W, Li XF, Jung IJ, Kim HJ. Repeated treatment with rituximab based on the assessment of peripheral circulating memory B cells in patients with relapsing neuromyelitis optica over 2 years. Arch Neurol. 2011 Nov;68(11):1412-20.
Pittock et al	Pittock SJ, Lennon VA, McKeon A, Mandrekar J, Weinshenker BG, Lucchinetti CF, O'Toole O, Wingerchuk DM. Eculizumab in AQP4-IgG-positive relapsing neuromyelitis optica spectrum disorders: an open-label pilot study. Lancet Neurol. 2013 Jun;12(6):554-62.

The approximate time (months) between the first 2 pretreatment relapses was visually enumerated from the published figures for each study (Figure 7-1, Figure 7-2, Figure 7-3, Figure 7-4, and Figure 7-5). Subjects with second pretreatment relapses occurring within 12 months of the first relapse were considered failures. Hazard rate (assuming exponential distribution) was calculated from each study based on this failure rate. Overall hazard rate for time to second pretreatment relapse was calculated based on the weighted average of individual hazard rates (weight was based on number of failures from each study). The findings from this analysis are summarized in Table 7-2.

Table 7-2 Estimated Pretreatment Hazard Rate for Time to Relapse

Study	Number of Subjects	No. of Subjects with Relapse by 1st Year	Percentage of Subjects Relapse Free by 1st Year	Hazard Rate (per year)
Bedi et al, 2011	23	14	39%	0.94
Costanzi et al, 2011	69	48	30.4%	1.2
Jacob et al, 2008	25	21	16%	1.83
Kim et al, Nov 2011	30	19	37%	0.99
Pittock et al, 2006	14	12	14.3%	1.94
Overall	161	114	29.2%	1.23

The weighted average of the hazard rates was determined as 1.33/year. As none of the cohorts in these studies received pure placebo treatment, the hazard rate for the placebo treatment for the seropositive subgroup in the current study is estimated to be 1.5/year. Conservatively, the hazard rate for the placebo treatment in seronegative subgroup is chosen as 1.0/year.

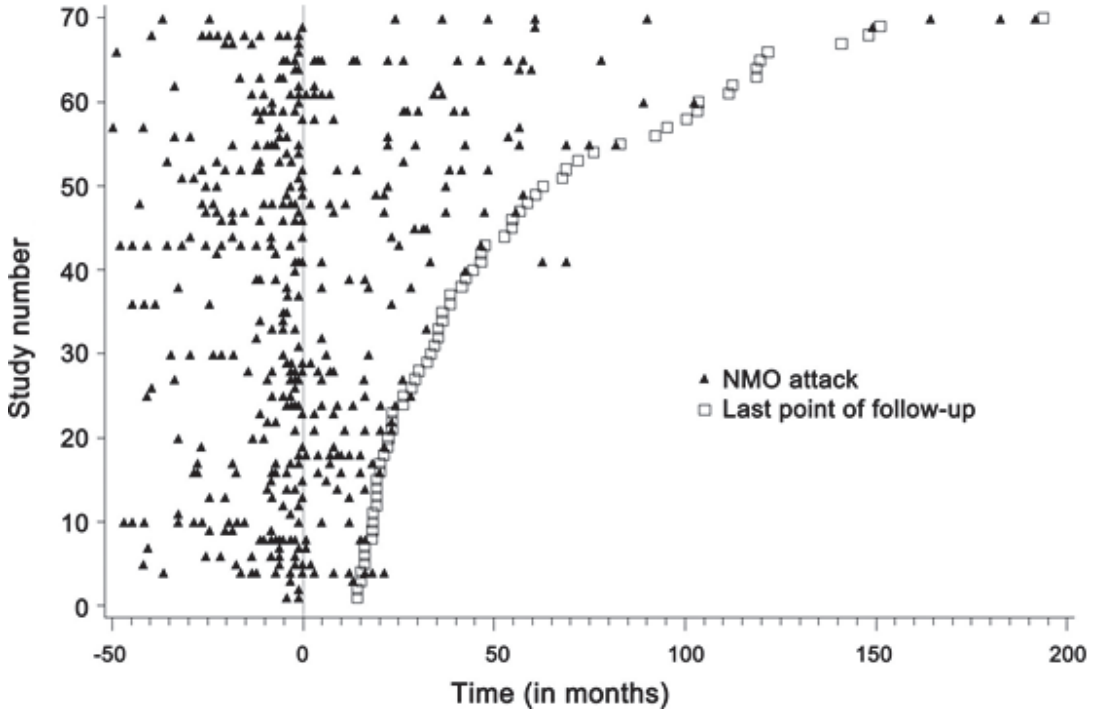

Figure 7-1 Neuromyelitis Optica Spectrum of Disorders Relapses Before and After Treatment with Azathioprine for 70 Patients with More Than 12 Months Follow Up

NMO = neuromyelitis optica.

Source: Constanzi et al, 2011.

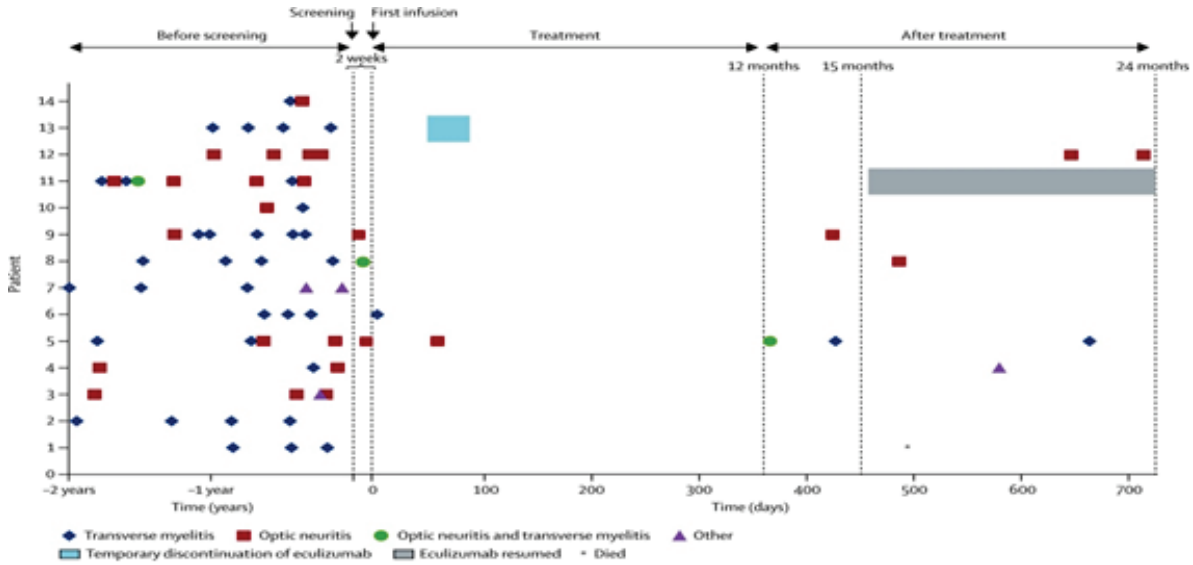

Figure 7-2 Attack Frequency Before, During, and After Eculizumab Treatment

Source: Pittock et al, 2006.

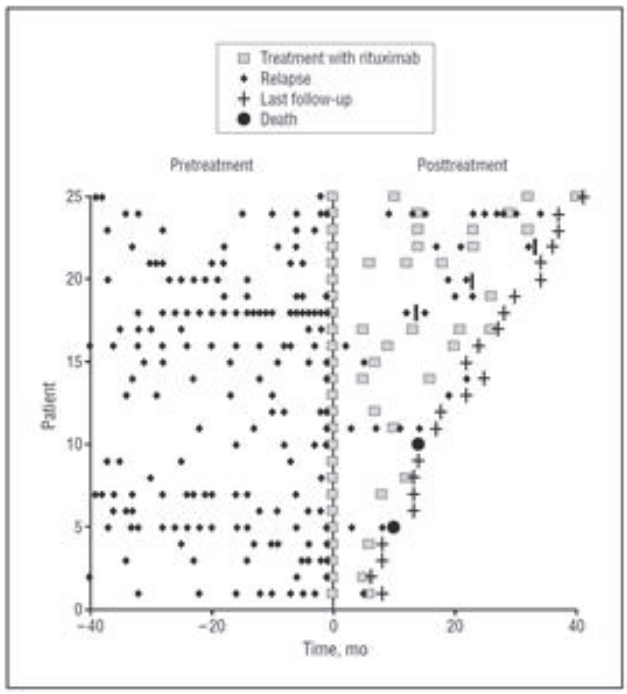

Figure 7-3 Relapses in Patients with Neuromyelitis Optica Before and After Treatment with Rituximab

mo = months.

Source: Jacob et al, 2008.

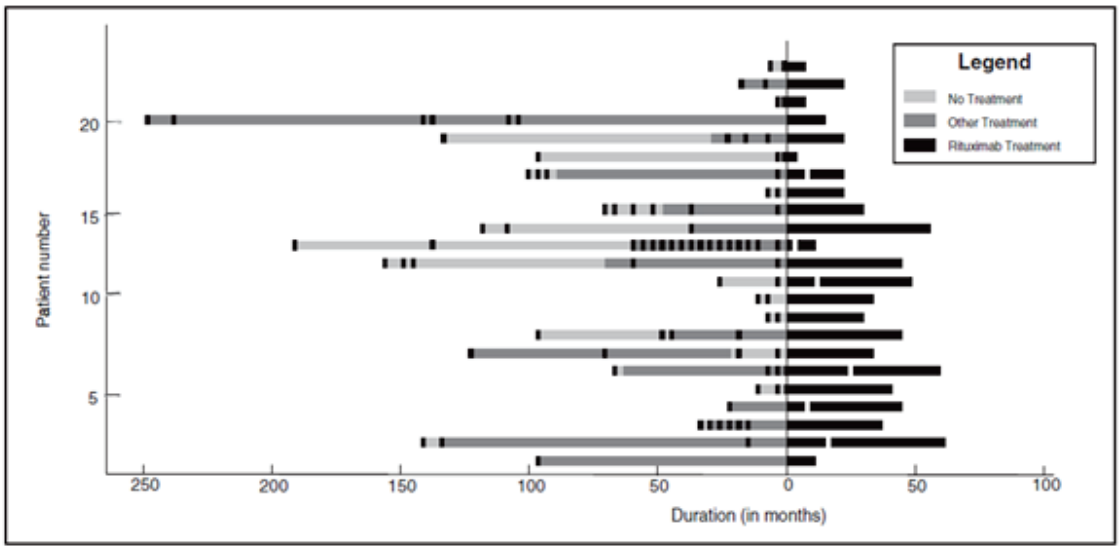

Figure 7-4 Disease Course of Patients Treated with Rituximab

Note: Disease course is depicted longitudinally, with each bar representing one patient. Each bar represents the type of treatment administered at a given time, a light gray bar representing no treatment, a dark gray representing other treatment (immunomodulatory, immunosuppressant) and black representing rituximab treatment. Each relapse is marked along the patient's course by a solid black or white mark. Time 0 on the X axis marks the initiation of rituximab treatment and divides pre- and post-rituximab treatment phases. Numbers on the X and Y axes represent months before/after rituximab initiation and patient number respectively. Most (17 of 23) patients had no recorded clinical relapses, and the remaining six had reduction in relapse frequency with rituximab treatment.

Source: Bedi et al, 2011.

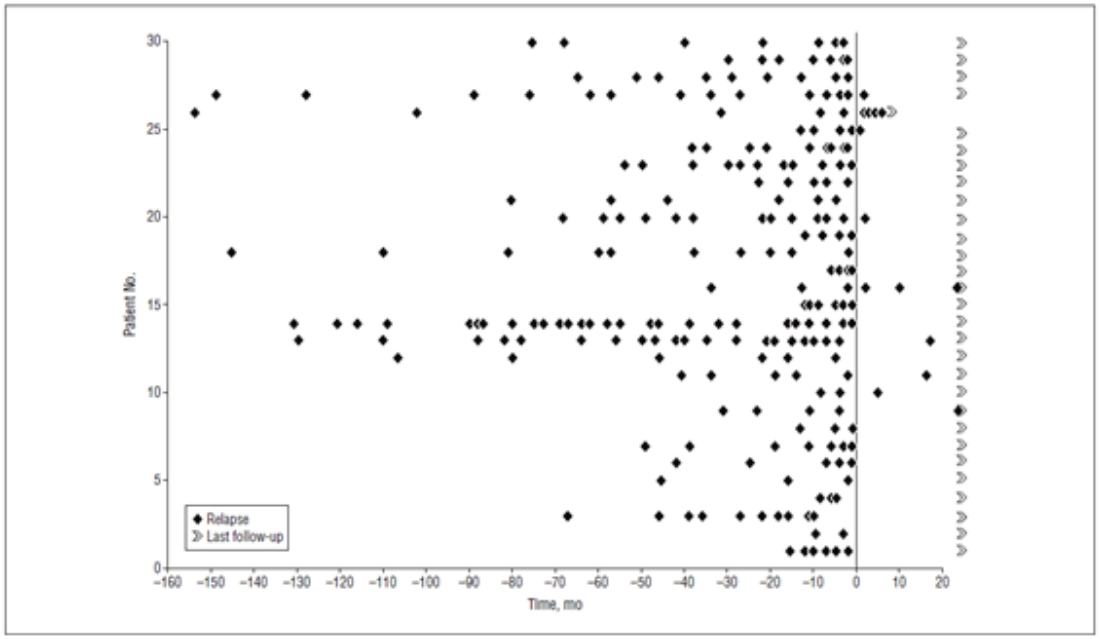

Figure 7-5 Relapses in Patients with Neuromyelitis Optica Before and After Treatment with Rituximab

mo = months.
Note: On the x-axis, 0 indicates start date of treatment. Each interrupted line on the y-axis represents a patient.
Source: Kim et al, Nov 2011.

The target treatment effect of 60% has been estimated by comparing similar empirical evidence from Costanzi et al, 2011 (treatment with azathioprine) and 3 articles from rituximab treatment (Bedi et al, 2011, Jacob et al, 2008, and Kim et al, Nov 2011). For each of these studies, the subjects who were relapse free within the first 12 months of respective treatment were visually identified (Table 7-3). Based on the survival rate thus obtained, the yearly constant hazard rate was calculated and the weighted average was determined.

Table 7-3 Calculated Yearly Constant Hazard Rate Based on Relapse-free Subjects in First 12 Months of Treatment

Study	N	No. of Subjects with Relapse by 1st Year	Percentage of Subjects Relapse-free by 1st Year	Hazard Rate (per year)
Bedi et al, 2011	23	4	83%	0.19
Jacob et al, 2008	25	7	72%	0.33
Kim et al, Nov 2011	30	6	80%	0.22
Overall	78	17	78%	0.25

The weighted average of the hazard for time to first relapse based on rituximab studies was determined as 0.258/year. Similarly, the hazard rate based on AZA treatment was calculated (Table 7-4). Based on this comparison, the relative reduction in risk of relapse was calculated as 57%.

Table 7-4 Relative Reduction in Risk of Relapse Based on Calculated Hazard Rate From Azathioprine and Rituximab NMO Studies

Agent	N	Events Within 1st Year	Estimated Hazard Rate	Hazard Ratio	Relative Reduction in Risk of Relapse
Rituximab	78	17	0.25	0.43	57%
Azathioprine	70	31	0.58		
Total	148	48	N/A		

N/A = not applicable; NMO = neuromyelitis optica

Since the treatment effect from a placebo-controlled study is not available at this point, the 60% relative reduction in hazard rate was used as a conservative estimate for the target treatment effect. It is recognized that there are a number of caveats behind this approach; such as, the cohorts compared were not randomized and the selection of the patient population from each study may not be a true reflection of the currently proposed study.

Appendix 2 Details of Futility Analysis

An unmasked interim analysis will be conducted for futility assessment by the IDMC when approximately 50% of the total planned AC-determined NMO/NMOSD attacks occur in this study. This will be triggered when approximately 34 AC-determined NMO/NMOSD attacks occur in the ITT population. The assessment of futility will be based on predictive power, which is based on the average conditional power calculated at the observed treatment effect at the time of the interim analysis. The study will be considered futile if the predictive power is < 20% in both the ITT population AND the AQP4-IgG seropositive cohort. Figure 7-6 provides estimated predictive power when observed treatment effect varies from 0 to 90% in relative reduction in risk of an AC-determined NMO/NMOSD attack in the ITT population. The 20% predictive power translates into a treatment effect of approximately 27%. A value of 50% and 90% predictive power will result in 43% and 60% treatment effects, respectively, which are minimally detectable treatment effect and target treatment effect for this study. The actual predictive power will be calculated based on observed data.

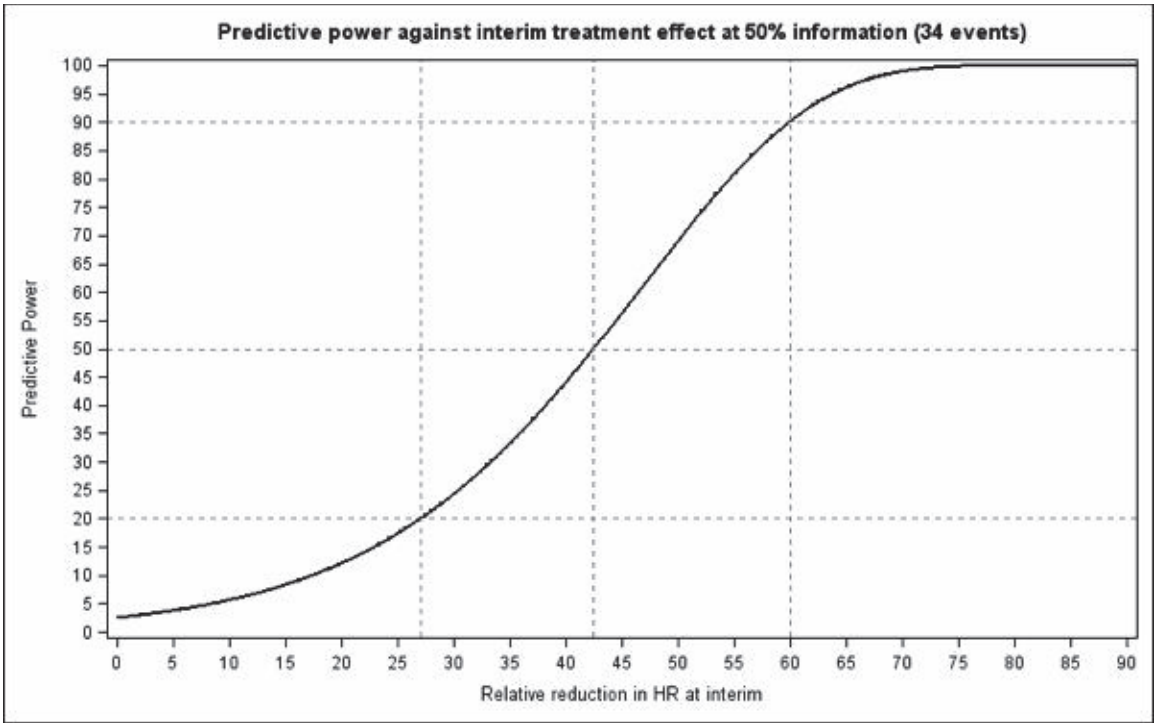

Figure 7-6 Estimated Predictive Power Against Observed Interim Treatment Effect When 50% of AC-determined NMO/NMOSD Attacks are Available

AC = Adjudication Committee; HR = hazard rate; NMO/NMOSD = neuromyelitis optica/neuromyelitis optica spectrum disorders

Appendix 3 Type I Error Control

To establish Type I error control, the primary endpoint, and the 4 secondary endpoints will be denoted by E1 through E5, respectively, as follows:

- E1: Time to AC-determined NMO/NMOSD attack
- E2: Worsening from baseline in EDSS at last visit during the RCP
- E3: Change from baseline in low-contrast visual acuity binocular score measured by low-contrast Landolt C Broken Rings Chart, at last visit during RCP
- E4: Cumulative total active MRI lesions (new Gd-enhancing or new/enlarging T2) during the RCP
- E5: Number of NMO/NMOSD-related in-patient hospitalizations

Based on 2 populations of interest (seropositive and ITT populations), this will result in testing 10 null hypotheses of no treatment effect labeled as follows (Table 7-5):

Table 7-5 List of Null Hypotheses Considered Under Multiplicity Adjustment Procedure

Hypothesis	Test
S1	Test E1 in seropositive subjects
S2	Test E2 in seropositive subjects
S3	Test E3 in seropositive subjects
S4	Test E4 in seropositive subjects
S5	Test E5 in seropositive subjects
O1	Test E1 in ITT population
O2	Test E2 in ITT population
O3	Test E3 in ITT population
O4	Test E4 in ITT population
O5	Test E5 in ITT population

The multiplicity adjustment strategy based on Bonferroni-based chain procedure (Bretz et al, 2009; Millen and Dmitrienko, 2011) for testing these 10 hypotheses is defined in Figure 7-7. Each hypothesis is represented by a rectangular box. The connections among the hypotheses are visualized using arrows. A solid arrow is used to define the decision path after a hypothesis is rejected, eg, the hypothesis O1 is tested if and only if the hypothesis S1 is rejected. The Bonferroni-based chain procedures are characterized by 2 rules:

- The alpha allocation rule specifies the initial distribution of the Type I error rate among the null hypotheses according to the relative importance of the null hypotheses

- The alpha propagation rule determines the process of re-distributing the available Type I error rate among the non-rejected null hypotheses after each rejection

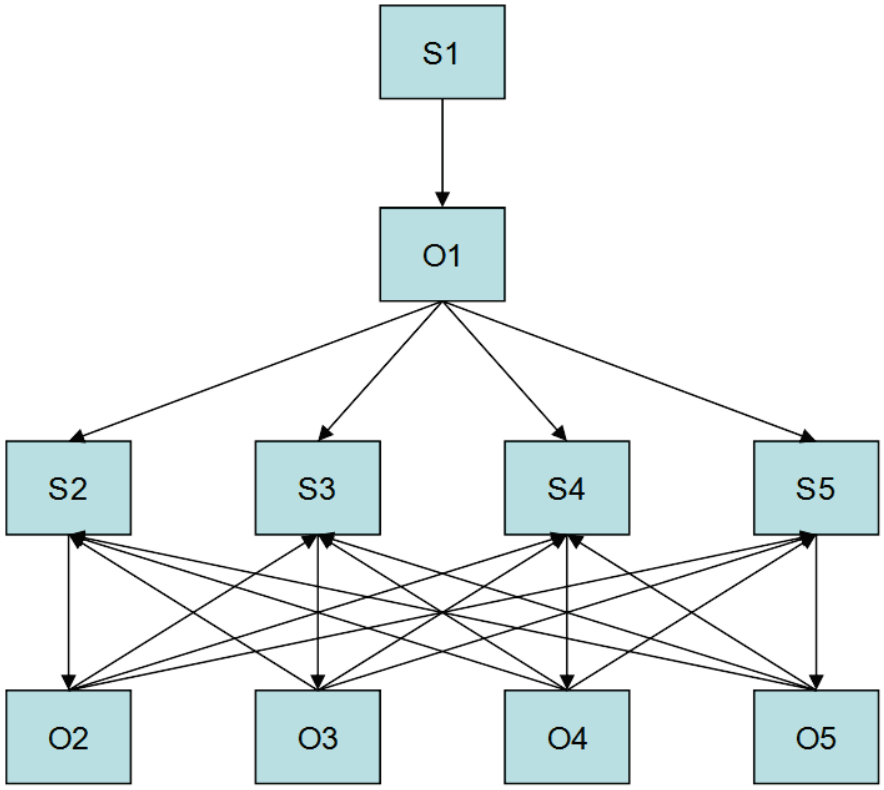

Figure 7-7 Multiplicity Adjustment Strategy

The alpha allocation rule is specified in Table 7-6. Specifically, the null hypothesis S1 receives the initial weight of 1 (ie, it is tested at the full $\alpha = 0.05$) and the other null hypotheses receive zero weights. The alpha propagation rule is specified in Table 7-7. The alpha propagation rules are defined using transition parameters, eg, the first row in Table 7-7 shows that the fraction of the Type I error rate used in testing the null hypothesis S1 will be transferred to the null hypothesis O1 if the null hypothesis S1 is rejected. Similarly, if the null hypothesis O1 is rejected, the available Type I error rate will be split equally among the null hypotheses S2, S3, S4, and S5 (a quarter of the available Type I error rate will be allocated to each null hypothesis). The alpha allocation and alpha propagation rule uniquely define the Bonferroni-based chain procedure and the associated multiplicity-adjusted p-values can be computed using Algorithm 2 given in Bretz et al (2009).

Table 7-6 Alpha Allocation Rule

Hypothesis	Initial Alpha Allocation
S1	α
S2	0
S3	0
S4	0
S5	0
O1	0
O2	0
O3	0
O4	0
O5	0

Table 7-7 Alpha Propagation Rule

Hypothesis	S1	S2	S3	S4	S5	O1	O2	O3	O4	O5
S1	0	0	0	0	0	1	0	0	0	0
S2	0	0	0	0	0	0	1	0	0	0
S3	0	0	0	0	0	0	0	1	0	0
S4	0	0	0	0	0	0	0	0	1	0
S5	0	0	0	0	0	0	0	0	0	1
O1	0	1/4	1/4	1/4	1/4	0	0	0	0	0
O2	0	0	1/3	1/3	1/3	0	0	0	0	0
O3	0	1/3	0	1/3	1/3	0	0	0	0	0
O4	0	1/3	1/3	0	1/3	0	0	0	0	0
O5	0	1/3	1/3	1/3	0	0	0	0	0	0

Appendix 4 Neuromyelitis Optica Attack Severity Score

Study CD-IA-MEDI-551-1155 is aimed at assessing the efficacy and safety of MEDI-551 in subjects with history of NMO/NMOSD who experienced at least one episode of attack. The primary endpoint is the time to an attack based on predefined criteria. The diagnosis of an attack, determined by the principal investigator at the investigational site will be reviewed and adjudicated by an AC. None of the endpoints in the study are based on the assessment of the severity of an attack. In discussion with experts in the field and with the FDA the significance of assessing the severity of NMO attack as exploratory analysis became apparent. There are no well-established and tested criteria to assess the severity of NMO/NMOSD attack. The criteria presented here are based on some data available in the literature (Pittock et al) and on an input from experts in the field. This scoring system is

exploratory in nature and its relevance and applicability will be assessed at the end of this study.

NMO/NMOSD attack severity score

A severity classification for each type of NMO/NMOSD attack (Optic Neuritis, Myelitis, Brain and Brainstem) is described in subsequent sections based on the respective subscale. Severity assessment (graded as 'minor' or 'major') will be based on a 'shift' in subscale scores corresponding to the domain of interest at the time of attack from the last assessment of the subscale.

When measured at the follow-up visit post an NMO/NMOSD attack, the 'shift' in the subscale score at the follow visit from the time of attack will be used to grade the 'recovery or improvement' as well (graded as 'minor' or 'major').

Subscale for each domain

Table 7-8 provides the subscales corresponding to each domain mentioned above. The V_A represents the corrected visual acuity assessment based on high-contrast Landolt Broken Ring C-chart.

Table 7-8 Subscale Score by Domain of NMO/NMOSD Attack

Domain	Subscale Score	Description
ON	0	Normal
	1	Scotoma[a] but $V_A \geq 50$ characters
	2	$V_A \geq$ 35-49 characters
	3	$V_A \geq$ 20-34 characters
	4	$V_A \geq$ 1-19 characters
	5	Counting fingers only
	6	Light perception only
	7	No light perception
Myelitis	0	Normal
	1	Abnormal signs (hyperreflexia, Babinski sign) without weakness
	2	Mild weakness (MRC grade 5- or 4+) in affected limb(s)
	3	Moderate weakness (grade 3 or 4) in 1 or 2 UMN muscles in affected limb(s)
	4	Moderate weakness (grade 3 or 4) in 3 UMN muscles in affected limb(s)
	5	Severe weakness (grade 2) in 1 or more muscles in affected limb(s)
	6	Some plegic (grade 0 or 1) muscles in 1 or more limbs
	7	Plegia (grade 0 or 1) of all muscles in 1 or more limbs
Brain	0	Normal

Table 7-8 Subscale Score by Domain of NMO/NMOSD Attack

Domain	Subscale Score	Description
	1	Drowsiness or mood changes only
	2	Mild confusion/disorientation (able to manage all self-care functions); mild focal impairment (mild aphasia, apraxia, agnosia, anorexia or drowsiness)
	3	Moderate confusion/disorientation (able to manage some self-care functions); moderate focal impairment (moderate aphasia, apraxia, agnosia, anorexia or drowsiness)
	4	Severe confusion/disorientation (unable to manage self-care functions); severe focal impairment (aphasia such that is unable to comprehend simple one step commands or speak 5 word sentences; severe apraxia, agnosia, anorexia or drowsiness)
	5	Stupor or coma
Brainstem	0	Normal
	1	Signs only (unsustained nystagmus, impaired saccadic pursuit, ocular dysmetria, mild facial weakness or sensory loss)
	2	Sustained conjugate nystagmus, incomplete INO, moderate facial weakness or sensory loss, or other mild disability; mild nausea and vomiting for 48 hours or longer without other explanation with vomiting not more than 3 times per day; intractable hiccups occurring more than 20 times per hour less than 6 hours per day
	3	Dyconjugate nystagmus (INO) or severe extraocular weakness, loss of facial sensation or facial paralysis (unilateral or bilateral), moderate dysarthria or dysphagia; moderate nausea and vomiting lasting 48 hours or longer without other explanation with vomiting between 3 and 7 times per day; intractable hiccups occurring more than 20 times per hour for 6 to 12 hours per day
	4	Severe dysarthria or dysphagia, almost complete ophthalmoplegia, or other severe disability of a cranial nerve/nerves; severe nausea and vomiting lasting 48 hours or longer without other explanation with vomiting occurring more than 7 times per day; intractable hiccups occurring more than 20 times per hour for more than 12 hours per day
	5	Inability to swallow or speak because of bulbar dysfunction; respiratory failure requiring intubation because of brainstem lesion.

[a] Scotoma assessment can be obtained from EDSS/FSS

Severity assessment

Based on the subscale score described in Table 7-8, the severity assessment of each type of NMO/NMOSD attack is provided in Table 7-9.

Table 7-9 NMO/NMOSD Attack Severity Grade

Domain	Subscale score at visit prior to attack	Subscale score at time of attack	Severity
ON, Myelitis, Brainstem	< 2	< 3	Minor
		≥ 3	Major
	≥ 2	Increase by 1 point	Minor
		Increase by ≥ 2 points	Major

Table 7-9 NMO/NMOSD Attack Severity Grade

Domain	Subscale score at visit prior to attack	Subscale score at time of attack	Severity
Brain	Not applicable	Increase by 1 point	Minor
		Increase by ≥ 2 points	Major
		Increase by 1 point	Minor
		Increase by ≥ 2 points	Major

Recovery assessment

Based on the subscale score described in Table 7-8, the recovery assessment of each type of NMO/NMOSD attack is provided in Table 7-10.

Table 7-10 NMO/NMOSD Attack Recovery Grade

Domain	Subscale score at time of attack	Improvement[a] at follow up visit	Recovery
ON, Myelitis, Brainstem	Any score	≤ 2	Minor
	≥ 3	> 2	Major
Brain	Any score	1	Minor
	≥ 2	>1	Major

[a] Improvement at follow up visit = Change from attack in subscale scores at follow up visit.

Appendix 5 WMWodds approach for mRS score analysis

Treatment effect based on mRS score during RCP will be evaluated by the WMWodds approach (Divine et al., Churilov et al.). This procedure estimates the odds that for two subjects chosen at random with one from each treatment group, the subject who received the investigational product will have a better outcome than the subject receiving standard treatment. Consider all possible pairs of observations where the first observation is taken from the MEDI-551 group (Y_T) and the second observation is taken from the placebo group (Y_C). If group Y_T included n observations and group Y_C included m observations, the two-group comparison would lead to the total of n × m pairs of observations. In each pair, the observation from the MEDI-551 arm will either be worse than the placebo observation, the same as the placebo observation, or better than the placebo observation. The probabilities that in a randomly chosen pair of MEDI-551 and placebo subjects, the MEDI-551 subject has worse outcome ($Prob(Y_T > Y_C)$), the same outcome ($Prob(Y_T = Y_C)$), or better outcome ($Prob(Y_T < Y_C)$), can then be estimated as the ratio of the number of pairs satisfying each of these individual conditions to the total number of n × m pairs. Given the possibility of tied observations between two treatment groups, the null hypothesis can be stated that the probability a subject under MEDI-551 treatment is worse than placebo is the same as the probability that a MEDI-551 treated subject is better than placebo accounting for ties, in other words, the probabilities $Prob(Y_T > Y_C) + 0.5Prob(Y_T = Y_C)$ and $Prob(Y_T < Y_C) + 0.5Prob(Y_T = Y_C)$ both being equal to 0.5. Thus, WMWodds test statistic is defined as follows:

$$\text{WMWodds} = \frac{Prob(Y_T < Y_C) + 0.5Prob(Y_T = Y_C)}{Prob(Y_T > Y_C) + 0.5Prob(Y_T = Y_C)}$$

Table 7-11 provides conceptual framework to identify number of pairs where mRS scores (5 and 6 combined together) in MEDI-551 group is same, better or worse than corresponding Placebo group scores. WMWodds is then defined by

$$= \frac{\text{Number of MEDI} - 551 \text{ better pairs} + 0.5 \times \text{Number of tied pairs}}{\text{Number of MEDI} - 551 \text{ worse pairs} + 0.5 \times \text{Number of tied pairs}}$$

The WMWodds and 95% CI will be calculated based on area under the curve (concordance index) from logistic regression.

Table 7-11 Conceptual outline of WMWodds approach

		Placebo					
	Score	0 (n=xx)	1 (n=xx)	2 (n=xx)	3 (n=xx)	4 (n=xx)	5/6 (n=xx)
MEDI-551	0 (n=xx)	Tied					
	1 (n=xx)		Tied	MEDI-551 is better than Placebo			
	2 (n=xx)			Tied			
	3 (n=xx)				Tied		
	4 (n=xx)	Placebo is better than MEDI-551				Tied	
	5/6 (n=xx)						Tied

Appendix 6 Kappa Statistic for Investigator and AC-determined Attack Agreement

Below is an investigation of the expected range of agreement between investigator and AC for NMO/NMOSD attack assessment for the current study. Figure 7-8 is based on 50,000 simulations. It is assumed that an AC-determined attack may be observed only if there is a site-determination of attack. X-axis presents the percentage of AC-determined attacks out of the total number of site determined attacks (positive percent agreement). A Kappa value of 1 (perfect agreement) is achieved when there are exactly 67 investigator-determined attacks which AC agrees in all cases. On the other extreme, when only 30% times AC agrees with the site, the Kappa value drops to less than 0.2 which is practically a high degree of statistical disagreement. The categories of Kappa value (according to Landis et al.) are shown on the right axis of the plot to determine the degree of agreement. It can be seen that based on the current design, one should expect as high as 75% of times AC should agree with site in order to get a Kappa value of 0.8 or higher. This also means that total number of investigator-determined attacks cannot be too high compared to target of 67 attacks, for example, higher than 90.

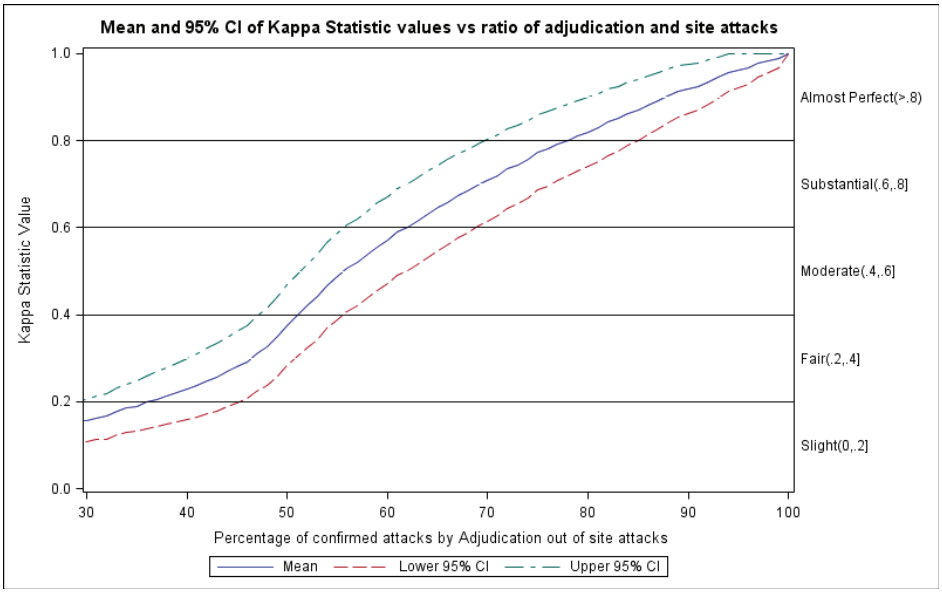

Figure 7-8 Agreement between Investigator and AC determination of Attacks

CI = Confidence interval