Official Title of Study:

A PHASE 3, RANDOMIZED, DOUBLE-BLIND CLINICAL TRIAL TO EVALUATE THE EFFICACY AND SAFETY OF ABATACEPT SC WITH STANDARD TREATMENT COMPARED TO STANDARD TREATMENT ALONE IN IMPROVING DISEASE ACTIVITY IN ADULTS WITH ACTIVE IDIOPATHIC INFLAMMATORY MYOPATHY (IIM)
PROTOCOL(S) IM101-611

NCT Number: NCT02971683

Document Date (Date in which document was last revised): December 15, 2020

STATISTICAL ANALYSIS PLAN
FOR CLINICAL STUDY REPORT

A PHASE 3, RANDOMIZED, DOUBLE-BLIND CLINICAL TRIAL TO EVALUATE THE EFFICACY AND SAFETY OF ABATACEPT SC WITH STANDARD TREATMENT COMPARED TO STANDARD TREATMENT ALONE IN IMPROVING DISEASE ACTIVITY IN ADULTS WITH ACTIVE IDIOPATHIC INFLAMMATORY MYOPATHY (IIM)

PROTOCOL(S) IM101-611

VERSION # 1.0

TABLE OF CONTENTS

LIST OF TABLES

1 BACKGROUND [redacted]

[redacted]

IIM who had persisting signs of active disease after treatment with glucocorticoids and one [redacted]

Research Hypothesis:

Subcutaneous abatacept (125mg weekly) in combination with standard treatment will achieve a higher rate of responders, defined as the percentage of subjects who achieve the IMACS DOI (International Myositis Assessment and Clinical Studies - definition of improvement), after 24 weeks of treatment versus standard treatment alone in adult subjects with active IIM.

Schedule of Analyses:

The analyses for this study is scheduled to be performed at 4 separate time points:

1) End of Double-Blind period (Week 24) for all subjects.
2) End of Open-Label period (Week 52) for all subjects.
3) End of Open-Label Extension period (Week 76) for sites in Japan. The purpose is to obtain 52 week abatacept efficacy data on all Japanese subjects for submission, specifically here the Japanese subjects originally randomized to placebo.
4) End of the study (completion of Long-Term Extension period and Follow-up period).

The Double-Blind period (Week 24) analysis will be performed once all subjects complete their Week 24 visit assessment or discontinue prematurely prior to Week 24. The Week 24 analysis will include the analyses of the primary efficacy endpoint, secondary and exploratory efficacy endpoints up to Week 24 and safety analyses. A subgroup analysis (selected outputs identified in DPP) on Japan subjects alone will be conducted for Japanese submission purposes. Actigraphy data and data from the Muscle Biopsy sub-study will be not be part of the initial Double-Blind period (Week 24) database lock and will be reported separately.

The Open-Label period (Week 52) analysis will be performed once all subjects complete 52 weeks of treatment in the study. A subgroup analysis (selected outputs identified in DPP) on Japan subjects alone will be conducted for Japanese submission purposes.

The Open-Label Extension period (Week 76) analysis will be performed, using only Japanese sites, once all Japanese subjects complete their Week 76 visit assessment. Selected outputs only identified in DPP.

The final analysis will be performed when all subjects have completed the 3 year long-term extension, (3.5 years for Japan subjects by including OLE data) and/or follow-up visits whichever is later.

This Statistical Analysis Plan provides details of the Week 24 analysis; details of the 3 other analyses are provided in APPENDIX 3.

In addition to these analyses, interim analyses are performed for the external, independent Data Monitoring Committee (DMC). The DMC was formed and charged with assessing the safety of Abatacept on a regular basis. The contents and methods of reports to the DMC are outlined in the DMC Charter[6].

2 STUDY DESCRIPTION

2.1 Study Design

This is a 24-week, Phase 3, randomized, double-blind, placebo-controlled, multicenter, study in subjects with active Idiopathic Inflammatory Myopathy (IIM; e.g., DM, PM, autoimmune necrotizing myopathy) on standard background treatment.

In addition to the 24-week double-blind period, the study includes:

- A 28-week open-label (OL) treatment period during which all subjects will receive SC abatacept with background treatment.
- An additional 24-week open-label of SC abatacept with continued therapy from the original OL period for subjects in Japan.
- A 3-year long-term extension of SC abatacept (156 weeks) with background treatment for all subjects except subjects in the U.S. and the Czech Republic.

The primary objective for this study is to compare the clinical efficacy of weekly abatacept in combination with standard treatment to standard treatment alone by assessing the percentage of subjects who achieve the IMACS DOI after 24 weeks of treatment versus standard treatment alone in adult subjects with active IIM.

Written informed consent must be provided by the subjects prior to undergoing any procedures during the screening phase.

Subjects who meet the study inclusion criteria enter into the 24-week Double-Blind (DB) period. During the DB period, approximately 150 subjects will be randomized in a 1:1 ratio to one of two treatment groups in a double-blinded manner as follows:

- Group A: Active Abatacept SC (125mg) weekly + standard treatment.
- Group B: Placebo Abatacept SC weekly + standard treatment.

At the time of randomization, subjects must be on a stable background treatment for IIM ("standard treatment" as defined in Section 3.4.1.1 of the protocol). No adjustment in background treatment will be allowed during the DB period.

During the 24-week DB period, subjects in either treatment group who meet the criteria for worsening (as defined in protocol Section 3.1.4) at any visit between Weeks 12 and 24 are permitted to use rescue therapy. The choice of rescue therapy is at the discretion of the investigator; however, rescue treatment may not be abatacept or a prohibited medication. Rescued subjects will remain blinded to study medication through Week 24 and will begin Open-Label (OL) abatacept treatment period after Week 24, however they will be considered not to have achieved the IMACS DOI for the purpose of the primary efficacy analysis.

The 24-week DB period is followed by a 28-week OL abatacept treatment period (OL period). During this period, subjects initially randomized to receive the combination of abatacept and standard treatment will continue this regimen. Subjects who are initially assigned to receive standard treatment alone will be switched to standard treatment plus abatacept. For the purposes of obtaining 1 year efficacy data in all Japan subjects (specifically for placebo randomized subjects) there will also be a 24-week open-label extension for subjects randomized in Japan. Background treatment may be adjusted (including the dose of corticosteroids and immunosuppressant's) at any time during the OL period. Use of immune globulin (intravenous [IVIG] or subcutaneous [SCIG]) is permitted during this period. Addition of any other prohibited medication is not permitted.

There will be a 3 year long-term extension period for subjects randomized in countries other than the U.S. and the Czech Republic. All subjects will receive standard treatment plus abatacept.

At the conclusion of the study, subjects who continue to demonstrate clinical benefit with study drug, i.e. abatacept, will be eligible to receive BMS supplied study drug based on local regulations.

In addition, subjects who discontinue treatment with study medication during the double-blind period will complete the Early Termination visit (which is the same visit as Day 169 visit). These subjects must also return for the double-blind period Day 169 post-randomization to complete the post-study drug efficacy visit. Subjects must also complete 2 follow-up visits (85 and 169 days after the last dose) during the 24-week post-treatment follow-up period, to perform safety and laboratory assessments. This period begins on the day of the Early Termination visit and applies only to subjects who do not start commercial abatacept therapy.

Subjects who discontinue treatment with abatacept during the open-label period or open-label extension for Japan will complete the Early Termination visit for the OL and 2 follow-up visits during the 24 week post-treatment follow-up period, to perform safety and laboratory assessments. If the subject receives treatment with an alternative source of abatacept, completing this post-treatment follow-up period is not required.

Subjects who discontinue treatment with abatacept during the long-term extension period will complete the Early Termination visit for the OL and two follow-up visits to perform safety and laboratory assessments. If the subject receives treatment with an alternative source of abatacept, completing this post-treatment follow-up period is not required.

2.2 Treatment Assignment

At the time of enrollment, immediately after written informed consent is obtained and before performing any study-related procedures, each subject will be assigned a unique sequential subject number beginning with 001, 002, 003, etc. for identification throughout the study by Interactive Recognition System (IxRS). This subject number must not be reused for any other participant in the study. The physician/coordinator must contact the Central Randomization System to enroll each subject into a centralized database at the beginning of the study (Screening Visit).

Randomization schedules were generated and kept by the Randomization Group within Drug Supply Management of Bristol-Myers Squibb. Each subject who is qualified for treatment is assigned a unique randomization number by IxRS. A randomization number will be assigned in the order in which subjects qualify for treatment, not in the order of study enrollment (screening visit).

Randomization will be stratified globally by 3 stratification variables: (1) Dermatomyositis (DM) vs. non-DM IIM, and (2) presence or absence of interstitial lung disease (ILD), (3) Japan vs. Rest of World (ROW).

Subjects are randomized in a 1:1 ratio within each of the strata to the abatacept or placebo treatment groups across all sites. Specific instructions for randomization into the Central Randomization System were provided in a separate manual.

Throughout the 24-week DB period of the study subject treatment assignment will remain blinded to clinical Investigational staff not involved in the preparation of the study medication.

2.3 Blinding and Unblinding

The subjects and clinical assessor(s) are not aware of which treatment is being administered to the subjects enrolled in the study. The pharmacist (or qualified drug preparation person) will be unblinded to study medication (SC Abatacept or SC placebo). The pharmacist (or qualified drug preparation person) know the scheduled assignments and would dispense the appropriate treatment accordingly.

Blinding is critical to the integrity of this clinical drug trial. However, in the event of a medical emergency or pregnancy in an individual subject, in which knowledge of the investigational product is critical to the subject's management, the blind for that subject may be broken by the treating physician.

Before breaking the blind of an individual subject's blinded treatment, the Investigator should determine that the information is necessary, i.e., that it will alter the subject's immediate management. In many cases, particularly when the emergency is clearly not investigational product-related, the problem may be properly managed by assuming that the subject is receiving active product without the need for unblinding.

BMS personnel will be blinded to treatment assignment until all subjects have completed the double-blind period, the database is locked for analysis (Week 24 analysis) and assessments of the primary efficacy endpoint are made.

2.4 Protocol Amendments

The table below summarizes the main purpose of each of the protocol amendments relevant to the analyses in this SAP. See the amendment for further details.

Table 2.4-1: Protocol Amendments

Document	Amendment Date	Main Purpose of Amendment
Amendment 02	10-Mar-2017	Clarify target disease population, exclusion criteria and rules for performing HRCT in at risk subjects.
Revised Protocol 03	09-May-2018	Clarify language, adjust criteria to improve subject selectio [redacted]

Table 2.4-1: Protocol Amendments

Document	Amendment Date	Main Purpose of Amendment
Revised Protocol 04	01-Feb-2019	Addition of descriptions for the Open Label Extension Period for Japan and the Long Term Extension Period.

3 OBJECTIVES

3.1 Primary

The primary objective for this study is to compare the clinical efficacy of weekly abatacept in combination with standard treatment to standard treatment alone by assessing the percentage of subjects who achieve the IMACS DOI by Week 24 compared to baseline, defined as:

- An improvement of ≥ 20% in 3 IMACS core measures, AND
- No more than 2 IMACS core measure scores worsen by ≥ 25%, AND
- Manual Muscle Test (MMT-8) may not decrease by ≥ 25%.

3.2 Secondary

1) To assess the clinical efficacy of weekly abatacept in combination with standard treatment to standard treatment alone by assessing the change in muscle endurance test using the Myositis Function Index (FI-2) from baseline to Week 24.
2) To assess the efficacy of abatacept in combination with standard treatment to standard treatment alone by assessing the mean change in functional disability using the Health Assessment Questionnaire-Disability Index (HAQ-DI) from baseline to Week 24.
3) To assess the efficacy of abatacept in combination with standard treatment to standard treatment alone by assessing the mean change in extra-muscular disease activity as defined by Myositis Disease Activity Assessment Tool (MDAAT) from baseline to Week 24.
4) To assess the efficacy of abatacept in combination with standard treatment to standard treatment alone in achieving improvement on the Myositis Response Criteria from baseline to Week 24.

[redacted]

[redacted]

4 ENDPOINTS

4.1 Primary Efficacy Endpoint

The primary endpoint for this study is the proportion of subjects who achieve IMACS DOI (as defined in Section 8.1.1) at Day 169 (Week 24) without rescue therapy.

4.2 Secondary Efficacy Endpoints

1) Mean change in muscle endurance using the myositis function index scores (FI-2) from baseline to Day 169 (Week 24).
2) Mean change in Health Assessment Questionnaire-Disability Index (HAQ-DI) from baseline to Day 169 (Week 24).
3) Mean change in Myositis Disease Activity Assessment Tool (MDAAT), based on the extramuscular disease activity VAS, from baseline to Day 169 (Week 24).
4) Myositis Response Criteria (MRC) total improvement score at Day 169 (Week 24).

4.3 Safety Endpoints

1) All adverse events (AEs/SAEs).
2) AEs of interest (serious infections, malignancies, injection site reactions, systemic reactions).
3) Laboratory test abnormalities.

[redacted]

[redacted]

[redacted]

5 SAMPLE SIZE AND POWER

A sample size of 150 subjects (75 subjects per treatment group) is planned based on the primary comparison of the proportion of subjects with IMACS DOI (as defined under the primary objective) at Day 169 (Week 24) between the SC abatacept and the SC placebo groups on a background of standard treatment.

A sample size of 150 subjects randomized in a 1:1 ratio to the SC abatacept group and SC placebo group will yield a power of approximately 90% to detect a treatment difference of 27% in the rate of IMACS DOI between the 2 treatment groups based on a continuity corrected chi-squared test. This power estimate assumes a two-sided alpha level of 5%, IMACS DOI rate of 60% in the abatacept group and 33% in the placebo group. The 60% estimate of IMACS DOI in the abatacept

group was based on reported rates of IMACS DOI in studies of patients with Myositis on combination therapy with another biologic agent. It is expected that the rate of IMACS DOI on abatacept in combination with standard treatment should be similar to that observed with the use of another biologic agent in combination with standard treatment. Table 5-1 below presents the power estimates for other considerations of proportion of subjects in IMAC DOI in the abatacept group assuming the total sample size of 150 subjects. Randomization will be stratified globally by 3 stratification variables: (1) disease diagnosis (Dermatomyositis [DM] vs. non-DM IIM), (2) Interstitial Lung Disease (ILD) (present vs. absent), (3) Japan vs. ROW.

Table 5-1: Estimates of Power for Various Assumed Rates of IMACS DOI at Week 24 Assuming Total sample Size of 150

Abatacept Rate	Placebo Rate	Delta	Power (%)
50%	20%	30%	96
	25%	25%	85
55%	25%	30%	95
	30%	25%	84
60%	25%	35%	~ 99
	30%	30%	95
	35%	25%	83

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

There are 6 analysis periods defined for the study: the 5 study periods ((i) pre-treatment period; (ii) 24-week DB period; (iii) the OL period; (iv) the OLE period for Japan; (v) the LTE period, incorporating the OLE period for Japan) and (vi) the cumulative abatacept period. These periods will be used in the safety summaries.

The pre-treatment period: covers the time period prior to the first dose date/time of study medication.

The 24-week Double-Blind period: starts at the first dose date/time of double-blind study medication and lasts until the last dose date of the double-blind period + 56 days if subjects discontinue the study during double-blind period, or until the first dose in the open-label period.

Since the dosing information for the Open-Label period, the Open-Label Extension period and the Long-Term Extension period is recorded on the same CRF page, it is not possible to derive the first/last dose for the different periods based on that CRF page. This is why start and end visit dates from other CRF pages will be used to distinguish between these 3 periods. Start and end visit dates are recorded for most periods, except the start of the OLE period for Japan subjects, this will be

assumed to be the end of the OL period. These dates will be used for deciding which dose belongs to which period.

The Open-Label period: starts at the first dose of the open-label period and lasts until the last dose date of open-label period + 56 if subjects discontinue the study during open-label period (or if subjects are from U.S. and Czech Republic sites), or until the first dose in the open-label extension period for subjects from Japanese sites, or the first dose in the long-term extension period for subjects from other sites.

The Open-Label Extension period (Japan subjects only): starts at the first dose of the open-label extension period (determined as the open-label period end visit date from the CRF) and lasts until the last dose date of open-label extension period + 56 if subjects discontinue the study during open-label extension period, or until the first dose in the long-term extension period.

The Long-Term Extension period (all subjects except for U.S. or Czech subjects): starts at the first dose of the long-term extension period and last until the last dose date of long-term extension period + 56.

The cumulative abatacept period: starts at the first dose date/time of abatacept and lasts until the last dose date of abatacept in the study + 56 days. For subjects randomized to placebo in the double-blind period the first dose date of abatacept will be the first dose date in the open-label period.

6.2 Treatment Regimens

There are 2 treatment groups in this study. Subjects will be randomized to either SC placebo or SC Abatacept (as randomized group) at the beginning of the study and each subject is expected to receive their randomized treatment (SC placebo or SC abatacept) throughout the double-blind period.

During the open-label period and beyond until study completion as per protocol/withdrawal all subjects will receive weekly SC abatacept injections.

6.3 Populations for Analyses

Efficacy endpoints will be summarized according to the Intent-to-Treat (ITT) population and safety endpoints will be summarized according to the as-treated analysis population. These populations are described briefly below along with other populations.

- Intent-to-Treat Analysis Population:

The Intent-to-Treat (ITT) analysis Population is defined to include all subjects randomized into the study for whom Case Report Form (CRF) data indicate that at least one dose of study medication (abatacept or placebo) is administered during the double-blind period.

All subjects who are randomized but never received study medication are to be excluded. Given the blinded nature of this study, it is reasonable to assume that subjects who discontinue prior to the receipt of study medication do so for reasons unrelated to study medication. Subjects are grouped according to the treatments to which they are randomized by IXRS.

ITT analysis population is also referred to as "All Randomized and Treated Subjects".

- As-Treated Analysis Population:

The As-Treated analysis population contains all subjects for whom CRF data indicate that at least one dose of study medication was administered during double-blind period.

Subjects are grouped according to the treatments to which they are randomized, except in cases where information is available, which indicates that a subject receives a different treatment for the entire course of double-blind period. In this case, the subject will be presented by the treatment they actually received.

The As-Treated analysis population will also be referred as "All Treated Subjects".

- Per-Protocol Analysis Population:

The Per-Protocol analysis population is defined as a subset of the ITT analysis population which excludes subjects who have relevant protocol deviations during the double-blind study period. Relevant protocol deviations are those protocol deviations that might affect the primary efficacy endpoint of the study. A programmable list of relevant protocol deviation criteria is included in APPENDIX 1.

The analysis of the primary efficacy endpoint will be repeated in the Per-Protocol analysis population. The Per-Protocol analysis population will exclude subjects with at least 1 relevant protocol deviation.

- Open-Label Treated Analysis Population:

The Open-Label treated analysis population will contain all subjects for whom CRF data indicate that at least one dose of study medication was administered during the open-label period. All Week 52 analyses will be performed using this population except otherwise specified.

- Japan Open-Label Extension Treated Analysis Population:

The Japan Open-Label Extension treated analysis population will contain all subjects from Japan for whom CRF data indicate that at least one dose of study medication was administered during the open-label extension period. All Week 76 analyses will be performed using this population except otherwise specified.

- Long-Term Extension Treated Analysis Population:

The Long-Term Extension treated analysis population will contain all subjects (except for U.S. or Czech subjects) for whom CRF data indicate that at least one dose of study medication was administered during the long-term extension period. All Year 3 (3.5 year in Japan subjects) analyses will be performed using this population except otherwise specified.

[redacted]

[redacted]

7 STATISTICAL ANALYSES

The analyses for this study is scheduled to be performed at 4 separate time points:

17) End of Double-Blind period (Week 24) for all subjects.
18) End of Open-Label period (Week 52) for all subjects.
19) End of Open-Label Extension period (Week 76) for sites in Japan.
20) And end of the study (completion of Long-Term Extension period and Follow-up period).

The double-blind period (Week 24) analysis will be performed once all subjects complete their Week 24 visit assessment or discontinue prematurely prior to Week 24. This analysis will include the analyses of the primary efficacy endpoint, secondary and exploratory efficacy endpoints up to Week 24 and safety analyses for the 24-week double-blind period. Safety analyses will include all the adverse events (AEs) reported in the study up to first dose date in the open-label period, or 56 days after the last dose date in the double-blind period whichever is earlier. Nominal p-values will be provided for the secondary endpoint measures at Day 169 (Week 24), no formal statistical testing will be conducted for any of the efficacy analyses related to the secondary endpoints at other time points and on any of exploratory endpoints. Appropriate summary statistics will be provided for the analysis of each endpoint.

The open-label period (Week 52) analysis will be performed once all subjects complete 52 weeks of treatment in the study. This analysis will include summaries of AEs reported during the open-label period. Efficacy endpoints will be summarized over time during the open-label period. No formal statistical testing will be conducted for any of the efficacy analyses. Appropriate summary statistics will be provided for the analysis of each endpoint. The 52-week analysis will be performed using the open-label treated analysis population. Selected outputs for the OL period will be detailed in the DPP.

The open-label extension period (Week 76) analysis will be performed, using only Japanese sites, once all subjects at Japanese sites complete 76 weeks of treatment in the study. This analysis will include summaries of AEs reported during the open-label extension period. Efficacy endpoints

will be summarized over time during the open-label extension period. No formal statistical testing will be conducted for any of the efficacy analyses. The open-label extension period analysis will be performed using the Japan open-label extension treated analysis population. Selected outputs for the OLE period will be detailed in the DPP.

The final analysis will be performed once all subjects complete the (3 year) long-term extension period and/or follow-up visits whichever is later. This analysis will include summaries of AEs reported during the long-term extension period (for open-label extension treated subjects) and the cumulative abatacept period (for abatacept treated subjects in the study) up to 56 days post last dose date in the study. The long-term extension period analysis will be performed using the long-term extension treated analysis population.

7.1 General Methods

The table below provides an overview of the efficacy analyses to be performed. These are thought to be robust in the case of the small marginal strata (Japan) frequencies of 20 or more (10 per treatment group); otherwise, some strata may be dropped from the models.

Table 7.1-1: Planned Efficacy Analyses

Measure of Interest	Analysis Method
IMACS DOI at Day 169 (Week 24)	**Primary Analysis:** Using a logistic regression model that includes treatment group, stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs. present), Region (Japan vs. ROW), baseline MMT8 score, baseline MDAAT, and baseline PGA. Point estimate of adjusted ORs, corresponding 95% CI and p-value will be provided. The proportion of subjects in IMACS DOI in each treatment group and corresponding 95% CI will also be provided.
Time to event (IMACS DOI at 2 consecutive time points) up to Day 169 (Weeks 12 and 24)	Point estimates and 95% CIs for the hazard ratio of abatacept versus placebo will be based on Cox proportional hazards models. The stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs. present), Region (Japan vs. ROW), baseline MMT8 score, baseline MDAAT and baseline PGA will be included in the model. Observations from subjects who discontinue due to lost to follow-up will be censored at the date of last available data in the study. The date of the first IMACS DOI of the 2 consecutive IMACS DOI events will be used in the analysis. Estimate the distribution of time to first occurrence of event using Kaplan-Meier method/curve.
Changes from baseline in continuous variables (i.e., FI-2 scores, MRC total improvement score, IMACS core components, [redacted]	Longitudinal (repeated measures) model [including treatment group, baseline value of variable and randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs. present), Region (Japan vs. ROW), time (categorical as study visit at which variable was measured), time by baseline value of variable and time by treatment as fixed effects and subject as a random effect], adjusted mean, SE, 95% CI for adjusted mean difference between treatment groups will be provided for all time

Table 7.1-1: Planned Efficacy Analyses

Measure of Interest	Analysis Method
	points. In addition, nominal p-values will be provided for the secondary endpoints of change from baseline at Day 169 (week 24) analysis.
Differences in proportions (i.e., proportion of subjects achieving IMACS DOI at Day 169 (Week 24), etc.)	Point estimate of response rate, 95% CI, point estimates and 95% CI of treatment difference adjusted for randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs. present), Region (Japan vs. ROW) (based on minimum risk weights).
Differences in proportions (i.e., proportion of subjects achieving IMACS DOI in myositis Ab negative or positive subjects, etc.)	Point estimate of response rate, 95% CI, point estimates and 95% CI of treatment difference adjusted for randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs. present), Region (Japan vs. ROW) (based on minimum risk weights).

All construction of CIs for differences in response rates between treatment groups will be based on minimum risk weights[7] to account for randomization stratification factors, unless otherwise noted. All construction of CIs for response rates within treatment group will be based on normal approximation methods, unless otherwise noted.

All construction of CIs for continuous measures will be based on a longitudinal repeated mixed effects model which includes treatment group, baseline value of variable, randomization stratification factors (diagnosis [DM vs. Non-DM], ILD [absent vs. present], Region [Japan vs. ROW]), time (as categorical variable), and time by treatment interaction as fixed effects and subject as a random effect. An unstructured covariance matrix will be used to model the correlation of the repeated measures within each subject. The SAS procedure PROC MIXED will be used with the restricted maximum likelihood (REML) method for estimation and the denominator degrees of freedom calculated according to the Kenward Roger method. The parameter estimations will be based on assumption of data being missing at random (MAR) using the method of restrictive maximum likelihood (REML). MAR means that given the observed data (e.g. for patients who discontinue up to the day they discontinue), the missingness mechanism does not depend on the unobserved data. In other words, this analysis assumes that a patient who drops out has the same conditional distribution (of the unobserved data after dropout given the observed data) as patients who share the same observed data and do not drop out. In case of non-convergence of the preferred longitudinal model or memory space issues the following back-up model will be used: a compound symmetry (CS) covariance matrix will be used to model the correlation of the repeated measures with each subject.

Unless otherwise specified, summary statistics will include the following:

- Summary statistics for continuous variables will include means, standard deviations, minima and maxima.

Summary statistics for qualitative or discrete variables will include the count and percentage.

7.2 Study Conduct

Relevant protocol deviations, which could have a major impact on the interpretation of the study results, will be identified prior to the unblinding of the database for all subjects who are randomized and receive double-blind medication.

Relevant protocol deviation criteria, including an assessment of the impact of any Covid-19 related deviations observed are listed in APPENDIX 1.

All subjects with relevant protocol deviations will be identified prior to the final database lock and unblinding of treatment assignment. All relevant protocol deviations will be listed and summarized by treatment group.

Following a case by case review of all incidences of Covid-19 related events (scheduling of visits, accurate receipt of medication) the impact of Covid-19 on the primary endpoint is considered minimal. The Per-Protocol analysis will exclude subjects with Covid-19 related deviations that could potentially impact the primary endpoint.

7.3 Study Population

Summary statistics of data pertaining to subject disposition, demographics, baseline characteristics, and medical history, including form of IIM, will be tabulated and displayed by randomized treatment group. No statistical test will be carried out for comparison of any baseline measurement between treatment groups.

These presentations will use the ITT analysis population and will be for "as-randomized" treatment regimen.

7.3.1 Subject Disposition

The disposition of subjects in the double-blind period will be presented as follows:

- Number of subjects enrolled at screening.
- Number of subjects randomized.
- Number of subjects treated in the double-blind period (i.e. who received at least 1 dose of study drug (abatacept or placebo).
- Number of subjects who discontinued during the double-blind period, together with reasons for discontinuation (taken from the eCRF status page).
- Number of subjects who completed the double-blind period, but who did not enter the open-label period and reason for not entering the open-label period (taken from the eCRF status page).

7.3.2 Demography and Baseline Disease Characteristics

Demographic and baseline disease characteristics will be summarized using the ITT analysis population. The summaries will be presented by as randomized treatment group and overall.

Continuous variables will be summarized using means and standard deviations, and categorical variables will be summarized using frequency distributions. The characteristics will be summarized based on the data captured on the CRF page, except for the randomization stratification which will be based on the IXRS data.

The following demographic and baseline characteristics will be summarized:

- Age.
- Age group (16-29, 30-39, 40-49, 50-59, ≥ 60 years).
- Gender.
- Race.
- Ethnicity.
- Geographic region:
  - North America: Canada and United States
  - South America: Argentina, Brazil, Chile, Columbia and Peru.
  - Europe: Italy, Romania, Russia, Spain and Turkey.
  - Asia: China, Hong Kong, Japan, Korea, and Taiwan.
  - Rest of World (ROW): Australia, Israel.
- Weight.
- Weight group (< 60, 60-100, > 100kg).
- Stratification factor Region (Japan vs. ROW).
- Stratification factor Disease diagnosis (DM vs. Non-DM).
- Stratification factor ILD (absent vs. present).
- Disease diagnosis: IIM type (DM, PM, autoimmune necrotizing myopathy).
- Randomization strata, Japan/Other, Disease Diagnosis; ILD:
  - Japan/Non-DM with ILD present.
  - Japan/Non-DM with ILD absent.
  - Japan/DM with ILD present.
  - Japan/ DM with ILD absent.
  - Non-Japan/Non-DM with ILD present.
  - Non-Japan/Non-DM with ILD absent.
  - Non-Japan/DM with ILD present.
  - Non-Japan/DM with ILD absent.
- Duration of disease in months.
- Duration of disease by disease diagnosis (DM vs. Non-DM).
- Medication History:
  - Number of prior medications used to treat Myositis: corticosteroids, immunosuppressant's, or biologics (1, 2, 3, > 3).
  - Number of prior DMARDS.

  - Number of current background Myositis medications: corticosteroids, immunosuppressant's (1, 2, 3, > 3).
- CK level at screening (continuous variable): number of subjects with elevated CK > 5 times the upper limit of normal at screening.
- Muscle enzyme levels (CK, AST, ALT, LDH, aldolase) at baseline (continuous variable).
- Baseline Physician's global assessment.
- Baseline Subject's global assessment.
- Baseline MMT-8.
- Baseline HAQ-DI (continuous variable).
- Baseline Myositis Function Index (FI-2) scores (continuous variables).
- Baseline Extramuscular Global Activity (as assessed by MDAAT).
- Baseline CDASI scores (continuous variables)
- Baseline MDI scores (continuous variables).

Any imbalances between treatment groups for the characteristics described above will be assessed when reviewing the summary tabulations and any differences deemed clinically relevant to the primary efficacy comparisons may be investigated by controlling for the characteristic in a supplemental analysis.

7.3.3 Medical History and Prior Medication

General medical history will be provided in a listing for the ITT analysis population.

Prior medications are defined as medications a subject started to take prior to the first SC injection of double-blind study medication. Prior medications will be summarized for each treatment group by drug class, generic name for the ITT analysis population.

Medications of special interest to Myositis treatment are the following categories: steroid and other steroid-sparing medications based on the WHO dictionary. A summary of prior medications of special interest will be provided by treatment group.

7.4 Extent of Exposure

7.4.1 Study Therapy

Because dosing for the open-label period, the open-label extension period (Japan) and the long-term extension period is collected on the same CRF, visit dates recorded on CRF pages that correspond to the first/last visits within these periods will be used to distinguish these periods and the exposure periods will be defined by visit start and end dates of visit periods recorded on the CRF.

7.4.1.1 Exposure during Double-Blind Period

Extent of exposure to double-blind study drug will be summarized using the as-treated analysis population in two ways:

- Number of injections.
- Number of days a subject is known to be on study drug.

The number of injections subjects received will be based on the study medication page of the CRF when an injection date is present.

The number of days the subject is known to be on study drug (exposure to study drug) during 24-week double-blind period is calculated as:

- For subjects who discontinued during 24-week double-blind period or did not enter the open-label period within 56 days of the last injection in the double-blind period:
- Exposure in days = (date of last injection in the 24-week double-blind period - date of first injection for double-blind period + 1) + 56.

For subjects who enter open-label period within 56 days of the last injection of 24-week double-blind period:

- Exposure in days = (date of first injection in the open-label period - date of first injection in the double-blind period).

The offset of 56 days is the length of 2 regular dosing cycles and represents approximately 4 times the half-life of abatacept in humans.

Summaries of exposure to study drug will show the distribution of the number of injections and days on drug, together with the means, standard deviations, medians, minima and maxima, by treatment group.

7.4.1.2 Exposure during Open-Label Period

The exposure to study drug during the open-label period will be calculated as follows:

Exposure to study drug in days for subjects who discontinued open-label period prior to the analysis cutoff date for the 24-week analysis is calculated as:

- Exposure in days = (date of last dose of study drug in the open-label period - date of first dose of study drug in open-label period + 1) + 56.

For subjects who are continuing beyond the analysis cutoff date for the 24-week analysis exposure is calculated as:

- Exposure in days = (data cutoff date - date of first dose of open-label period + 1).

The data cutoff date is the date of the last data change in the database prior to the DB lock.

7.4.1.3 Exposure during Open-Label Extension Period

The exposure to study drug during the open-label extension period will be calculated as follows:

Exposure to study drug in days for subjects who discontinued open-label extension period prior to the analysis cutoff date for the 28-week analysis is calculated as:

- Exposure in days = (date of last dose of study drug in the open-label extension period - date of first dose of study drug in open-label extension period + 1) + 56.

For subjects who are continuing beyond the analysis cutoff date for the 28-week analysis exposure is calculated as:

- Exposure in days = (data cutoff date - date of first dose of open-label extension period + 1).

The start date of the OLE period is not recorded for Japan subjects but is assumed to be the end date of the OL period.

7.4.1.4 Exposure during Long-Term Extension Period

The exposure to study drug during the long-term extension period will be calculated as follows:

Exposure to study drug in days for subjects who discontinued long-term extension period prior to the analysis cutoff date for the 24-week analysis is calculated as:

- Exposure in days = (date of last dose of study drug in the long-term extension period - date of first dose of study drug in long-term extension period + 1) + 56.

For subjects who are continuing beyond the analysis cutoff date for the 24-week analysis exposure is calculated as:

- Exposure in days = (data cutoff date - date of first dose of long-term extension period + 1).

7.4.2 Discontinuation of Study Therapy

Discontinuation from study medication is defined as subject's termination of the study medication (abatacept or placebo) without resumption prior to study completion. A subject will be considered as discontinued from study period if and only if the subject status CRF page indicates he/she did not complete either of the study periods. A separate subject status CRF page is completed at discontinuation or end of (1) double-blind period, (2) open-label period, (3) open-label extension period and (4) long-term extension period.

The number of subjects who discontinue during each of the 4 analysis periods; together with the reasons for discontinuation will be summarized by the randomized group as described in Section 7.3.1. The reasons for discontinuation will be taken from the subject status pages of the CRF for each period.

7.4.3 Treatment Compliance

The number of subjects with missed injections during the double-blind period (excluding missed injections due to premature discontinuation from study) will be summarized by treatment group and number of missed injections. This summary will be performed with the as-treated analysis population. A separate summary will also be presented for the other periods.

7.4.4 Disease Worsening and Rescue Therapy

The number of subjects who meet the criteria for disease worsening (Section 3.1.5 of the Protocol) during the DB period will be summarized by treatment group. For subjects who meet worsening disease criteria, the mean day of the first visit meeting the criteria will be calculated by treatment group. The number of subjects meeting the criteria for disease worsening who undergo rescue therapy will be summarized by treatment group. For subjects who undergo rescue therapy, the mean day of start of rescue therapy will be calculated by treatment group. The changes to concomitant therapy used as rescue therapy will be summarized (increase in dose of current therapy, change in therapy and/or addition of therapy).

[redacted]

7.5 Efficacy

The primary and secondary efficacy analyses below will be performed using the IIT analysis population by as randomized treatment group. The primary analysis will be repeated using the PP analysis population. All other efficacy analyses will be performed using the ITT analysis population, unless otherwise specified.

For all time to event analyses, unless otherwise stated, subjects who do not achieve an efficacy endpoint (e.g., response) will be censored at the time of last non-missing visit in the double-blind period.

7.5.1 Hierarchical Testing for Primary and Key Secondary Endpoints

If the test for the primary endpoint (proportion of subjects who achieved IMACS DOI at Day 169) is statistically significant at a level of 5%, then a hierarchical approach for statistical testing will be used for the secondary endpoints. Hierarchical ordering of the secondary endpoints is as follows:

- Mean change from baseline in HAQ-DI at Day 169.
- MRC total improvement score at Day 169.
- Mean change from baseline in muscle endurance using FI-2 at Day 169.
- Mean change from baseline in MDAAT, based on the extramuscular disease activity VAS, at Day 169.

P-values will be presented for each of these endpoints. However, endpoints should not be interpreted for significance if they have a rank lower than that endpoint whose null hypothesis was the first that could not be rejected. That is, the significance of endpoints should not be interpreted if they have a rank lower than that endpoint whose test was the first to be non-statistically significant (p-value > 0.05).

7.5.2 Primary Efficacy Analysis

The primary comparison of proportion of subjects in IMACS DOI at Week 24 between the SC abatacept and placebo groups will be assessed using a logistic regression model. To account for randomization stratification, the logistic regression model will adjust for the randomization stratification variables of disease diagnosis (DM vs. Non-DM), Interstitial Lung disease status (present vs. absent), and Japan vs. ROW; baseline MMT8 score, baseline MDAAT, baseline PGA and treatment group will also be included in the model. Point estimates of the adjusted Odds Ratios (OR) of the odds of achieving IMACS DOI in the abatacept group compared to the placebo group, corresponding 95% CI and p-value will be provided. Point estimates of the proportion of subjects in IMACS DOI in each treatment group and corresponding 95% CI will also be provided. All subjects who discontinue study medication prematurely prior to reaching the Week 24 assessment visit and/or received rescued medication at any time during the 24 week double-blind period, will be considered as not achieving IMACS DOI for the primary analysis. Efficacy assessments collected at Week 24 on subjects who discontinue study medication prior to Week 24 will not be included in the primary analysis of the primary endpoint, but will be included in a sensitivity analysis of the primary endpoint. Details of imputation rules for classification of IMACS DOI response for subjects with missing components of IMACS are provided in Section 8.1.9. A table summarizing the number of imputed values for the primary endpoint by treatment group and missing data pattern will be provided. A summary of number of missing IMACS core measure item by treatment group will also be provided for subjects whose IMACS data are missing at Week 24 because of missing component data.

The primary estimand of the treatment difference in the proportion of subjects with IMACS DOI at Day 169 is the odds ratio (abatacept versus placebo) in the ITT analysis population that would be observed if all subjects who prematurely discontinued or received rescue medication before Day 169 or had missing IMACS DOI data would be non-responders.

7.5.3 Sensitivity Analyses on the Primary Endpoint

7.5.3.1 Assess Impact of Rescue Medication

Sensitivity analyses of the primary endpoint will be performed based on the following scenarios:

- All subjects who meet criteria for disease worsening will be imputed as non-responders (i.e. not meeting IMACS DOI) whether they used rescue medication or not.
- All subjects who received rescue medication but continued in the double-blind period to Day 169 will be assessed for IMACS DOI based on their disease assessment on Day 169.

These sensitivity analyses will help assess the impact of the use of rescue medication on the primary endpoint.

7.5.3.2 Tipping Point Analyses

The robustness of the results of the primary endpoint will be assessed by varying the assumption of the missing data at the Week 24 (Day 169) based on the tipping point approach. This analysis will challenge the assumption that the effect of abatacept and placebo is the same for missing data by varying the percentage of missing considered as non-responders in both the abatacept and the placebo treatment group.

The missing data will be imputed using multiple imputation assuming the data are MNAR. The multiple imputation will utilize a logistic regression model that includes treatment group, age, gender, stratification variables of disease diagnosis (DM vs. Non-DM), Interstitial Lung disease status (present vs. absent), and Japan vs. ROW; baseline PGA score, baseline MMT-8 score and baseline MDAAT. The imputation will consider various scenarios including where dropouts on abatacept have lower response rate than dropouts on placebo. A series of analyses will be performed varying the distribution of responses among the missing data in each treatment group until the analysis conclusion of a statistically significant treatment effect no longer holds. The combination of number of responses among the missing data in each treatment group that overturns the primary results will represent a tipping point. An interpretation of clinical plausibility of the assumption underlying the tipping point will be provided.

Imputed data will consist of 100 imputed datasets; for each imputed dataset the missing IMACS DOI response will be imputed using the LOGISITIC method with a random seed number of 123611. Each of the imputed datasets will be analyzed using a similar logistic regression model used for the primary endpoint to get the estimate of odds ratio and 95% CI. Results from the analyses of each imputed datasets ORs and corresponding 95% CI will be combined using Rubin's imputation rules (using SAS MIANALYZE procedure) to produce the pooled estimate of OR and corresponding 95%, a pooled p-value for the test of null hypothesis of no treatment effect (i.e. OR = 1).

Analyses will be conducted by varying the log odds of achieving an IMACS DOI response (i.e. LOGIT of the probability of IMACS DOI response) among the missing data within each treatment arm. Resulting p-values will be displayed in a matrix that presents a range of logits corresponding to varying probabilities of response for missing values in both treatment arms; starting with a logit corresponding to a probability of response = 0.14 (this assumes 1 response out of 7 missing data) with increments of 0.3 on the logit until the logit is reached corresponding to a probability of response of .86 (this assumes 6 responses out of 7 missing data). The tipping point can then be identified (i.e. where a statistically significant treatment effect no longer holds). The adjustment in the logit function will be specified using the statement MNAR ADJUST available in SAS 9.3 or higher.

7.5.3.3 Sensitivity Analyses Based on Day 169 Data Collected after Discontinuation from Study

Subjects who discontinued the study prior to the Week 24 (Day 169) assessment are required by protocol to return for their Week 24 study visit similar to subjects continuing in the study; in

addition to the tipping point analysis above, the data collected on the discontinued subjects who did not receive rescue medication will be used to perform an additional sensitivity analysis of the primary endpoint. A similar model used for the primary analysis will be used; the logistic regression model will include randomization stratification variables of disease diagnosis (DM vs. Non-DM), Interstitial Lung disease status (present vs. absent), and Japan vs. ROW; baseline MMT-8 score, baseline MDAAT, baseline PGA and treatment group. Point estimate of the adjusted Odds Ratios (OR) of the odds of achieving IMACS DOI in the abatacept group compared to the placebo group, corresponding 95% CI will be provided, however no p-value will be calculated. Unadjusted point estimate of the proportion of subjects in IMACS DOI in each treatment group and corresponding 95% CI will also be provided. Subjects with missing Week 24 assessment will be considered non-responders in this sensitivity analysis.

7.5.4 Secondary Efficacy Analysis

In addition to the primary efficacy endpoint, analyses related to the secondary [redacted] will also be included in the Week 24 analysis. The secondary efficacy assessment of change from baseline in muscle endurance using FI-2 up to Day 169 (Week 24) will be assessed using a longitudinal (repeated measures) model including treatment group, baseline FI-2 value and randomization stratification factors: diagnosis (DM vs. Non-DM), ILD, Japan vs. ROW, time (categorical as study visit at which FI-2 is measured), time by baseline FI-2 and time by treatment interactions as fixed effects and subject as a random effect. Adjusted mean, SE, 95% CI for adjusted mean difference between treatment groups by study visit will be provided.

A similar longitudinal model will also be used to assess the change from baseline in HAQ-DI up to Day 169 (Week 24). The model will include treatment group, baseline value HAQ-DI value, randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD, Region (Japan vs. ROW), time (categorical as study visit at which HAQ-DI is measured), time by baseline HAQ-DI and time by treatment interactions as fixed effects and subject as a random effect. Adjusted mean, SE, 95% CI for adjusted mean difference between treatment group by study visit will also be provided.

The secondary efficacy assessment of change from baseline in extra-muscular disease activity MDAAT value will also be assessed using a longitudinal (repeated measures) model including treatment group, baseline value extra-muscular disease activity MDAAT value, randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD, Region (Japan vs. ROW), time (categorical as study visit at which extra-muscular disease activity MDAAT is measured), time by baseline extra-muscular disease activity MDAAT and time by treatment interactions as fixed effects and subject as a random effect. Adjusted mean, SE, 95% CI for adjusted mean difference between treatment groups by study visit will also be provided.

A similar longitudinal model will also be used to assess the Myositis Response Criteria (MRC) total improvement score up to Day 169 (Week 24). The mixed model will include treatment group and randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs. present), Region (Japan vs. ROW), time (categorical as study visit at which MRC is measured)

and time by treatment interactions as fixed effects and subject as a random effect. Adjusted mean, SE, 95% CI for adjusted mean difference between treatment groups by study visit will be provided.

Nominal p-values will be provided for the secondary efficacy endpoints at Day 169 i.e. mean change from baseline at Day 169 for each of the secondary endpoints: FI-2, HAQ-DI, extra-muscular disease activity MDAAT and MRC. The mixed longitudinal model will account for any missing values on any of the secondary efficacy endpoints.

The primary analyses of each of the secondary endpoints discussed above will be based on the data observed while subjects are on treatment in the study; a sensitivity analysis will be performed separately for each secondary endpoint that includes all the data collected at Week 24 on subjects who discontinue study medication prior to Week 24. Similar models as used for the primary analysis of these endpoints will be applied for the sensitivity analysis.

[redacted]

[redacted]

[redacted]

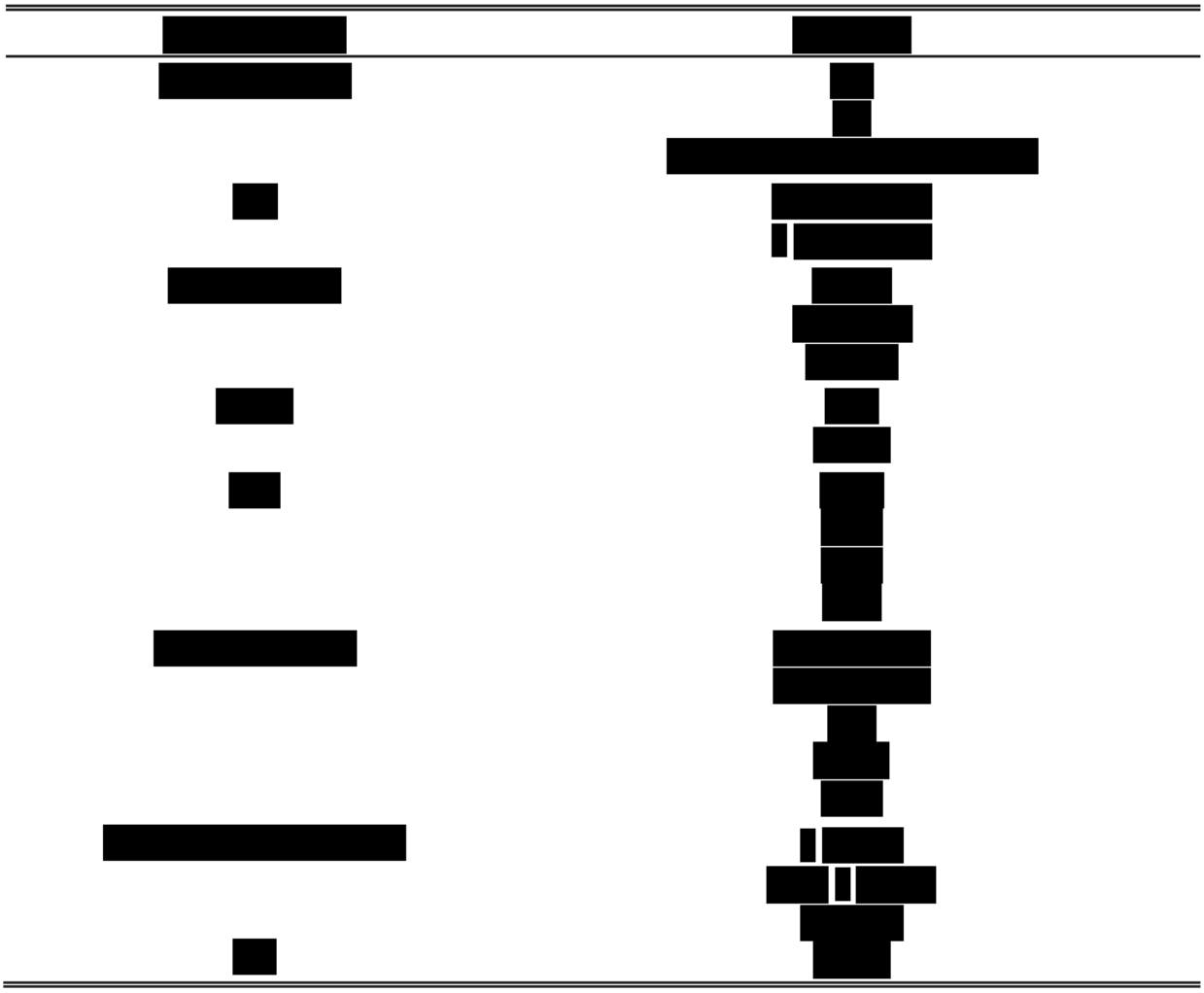
[redacted]

Non-DM), ILD (absent vs. present), Region (Japan vs. ROW), time (categorical as study visits at

[redacted]

central lab for processing and undergo pathological review and scoring (see Muscle Biopsy [redacted]

[redacted]



[redacted]

7.6 Safety

Analysis of all safety data will follow the BMS standard safety data conventions and supplements to the standard conventions for the abatacept programs[8].

The evaluation of drug safety is based primarily on clinical AEs, vital signs and laboratory abnormalities reported during the double-blind period. All safety presentations will be based on the as-treated analysis population by treatment group in the double-blind period.

Frequency distributions and individual listings of all AEs will be generated. Changes in clinical laboratory test results from baseline will be summarized. Laboratory marked abnormalities using pre-defined abnormality criteria will also be descriptively summarized. There will be no statistical testing of group difference with respect to frequencies of laboratory marked abnormalities.

7.6.1 Adverse Events

Adverse events are recorded by the investigators on the Serious and Non-Serious Adverse Event page(s) of the CRF. All investigators are required to report the nature, the onset and resolution date, intensity, action taken, treatment required for event, and to express their opinion regarding the relationship between the AE and the study medication.

All AEs are coded and grouped into Preferred Terms (PT) by System Organ Class (SOC), using the most current version of the Medical Dictionary for Regulatory Activities (MedDRA), at the time of data base lock. Listings and summaries will be based on the resulting SOCs and PTs.

All reported AEs and SAEs will be listed, indicating the subject id, treatment group, age, gender, race, day of onset relative to start of dosing, resolution date, investigator-assessment of relationship to study drug, investigator-assessment of intensity of event, action taken regarding study drug and whether treatment was required for the event.

All adverse events with an outcome of death will be listed.

Any AEs that occur before the first injection of double-blind study drug or more than 56 days after the last injection during the double-blind period will be excluded from safety summaries in the Week 24 analysis. However, all reported AEs, including AEs that occurred more than 56 days post last dose for subjects who discontinued the study during the double-blind period or completed the double-blind period but chose not to enter the open-label period will appear in the above mentioned comprehensive listing.

Summary information (the number and percent of subjects) regarding AEs (for serious and non-serious events) will be tabulated by SOC, PT and treatment group for:

- Serious adverse events (SAE).
- SAEs related to study drug.
- AEs leading to study drug discontinuation.
- AEs related to study drug.
- AEs related to study medication categorized by intensity.
- Most common related AEs (reported in at least 2% of subjects in any treatment group).
- AEs by intensity.
- All AEs (serious or non-serious).
- Most common (serious or non-serious) AEs (reported in at least 5% of subjects in any treatment group).
- Laboratory AEs are laboratory results identified by the Investigator as AEs and thus reported on the AE pages of the CRF. Any such AE will be included in the respective AE summaries.

7.6.2 Adverse Events of Interest

For all adverse events of special interest, treatment groups will be compared descriptively using risk differences and corresponding 95% CI and presented graphically using a forest plot in decreasing order of risk difference.

7.6.2.1 Infections

All reported infections and infestations within the SOC: Infections and infestations occurring during the double-blind period are included in the summaries described in Section 7.6.1.

7.6.2.2 Malignancy

All events in the MedDRA Maintenance and Support Services Organization (MSSO) malignancies Structured MedDRA Query (SMQ) list occurring during the double-blind period will be captured and reported. These events will be summarized by preferred term and treatment group and will also be listed.

7.6.2.3 Autoimmune Disorders

The frequency of pre-specified autoimmune disorders, defined using a BMS custom Autoimmune Disorder SMQ, occurring during the double-blind period will be provided. Autoimmune disorders occurring during the double-blind will also be summarized by intensity. All reported autoimmune disorders occurring during the double-blind period will be listed.

7.6.2.4 Injection Adverse Events

A. Systemic Injection Reactions are defined pre-specified AEs (such as hypersensitivity reactions) occurring during the first 24 hours after SC injection. The pre-specified list of AEs is defined using a BMS custom SMQ for abatacept SC program.

B. Injection Site Reaction Adverse Events: Summaries of injection site reactions will be done in two ways: using the pre-specified list of injection site reactions (cSMQ) and using the standard HLT lists for Injection site reactions, Administration site reactions NEC, and Application and instillation site reactions.

The number and percent of subjects experiencing the injection reactions defined in A and B above will be summarized by SOC and PT and treatment group. The distribution of the injection reaction events of interest by severity will also be provided. Serious events of interest during the double-blind period will be summarized by SOC, PT and treatment group. Most frequent (reported in 5% of subjects or more in any treatment group) AEs will also be provided. All reported injection reactions events of interest occurring during the double-blind period will be listed.

In addition, a pre-specified local injection site reaction event of interest, defined as the AEs with the following PT that occur at the site of SC injection: *Erythema, Pain, Pruritus, Haematoma, and Swelling* will be summarized by severity.

7.6.3 Laboratory Data

Safety summaries based on laboratory test results include evaluation of laboratory abnormalities and changes in laboratory test values from baseline (baseline addressed further in Section 8.2). For summaries of the double-blind period, laboratory data obtained during treatment or up to 56 days after last injection of study medication in the double-blind period but no later than the start of the open-label period dosing will be eligible for analyses.

Laboratory abnormalities are identified using a pre-defined set of marked abnormality (MA) criteria (listed in Safety Data Convention). The frequency of subjects with laboratory MAs during the double-blind period based on pre-specified criteria will be tabulated by treatment group, for each analyte.

Laboratory measurements and their changes from baseline will be summarized by nominal visit and treatment group, for protocol specified analytes. Subjects who have laboratory measures at baseline and corresponding measures on at least one scheduled visit following administration of study drug will be included in the laboratory analyte assessment. Note that not all subjects have laboratory determinations for all analytes at all visits, and therefore the sample size may vary from analyte to analyte at each time point.

In addition shift tables for the following laboratory tests will be provided. The categorical classifications of these tests will be determined based on the cut points specified below and cross-tabulation of classification between baseline and each of the scheduled visits will be presented by treatment groups.

Table 7.6.3-1: Definitions of Analyte Cut Points

Analyte (Test Code)	Cut Points
Platelet Count (PLAT)	$< 100x10^9$ c/L
	$\geq 100 x10^9$ c/L
Alanine Aminotransferase (ALT)	< 3xULN
	3 - ≤ 5xULN
	> 5xULN
Aspartate Aminotransferase (AST)	< 3xULN
	3 - ≤ 5xULN
	> 5xULN
Neutrophils (absolute)	$< 0.5x10^9$ c/L
	$0.5 - \leq 15x10^9$ c/L
	$> 15x10^9$ c/L
G-Glutamyl Transferase (GGT)	< 3xULN
	3 - ≤ 5xULN
	> 5xULN

7.6.4 Vital Signs

- Summary statistics for vital sign measurements (systolic BP, diastolic BP, heart rate and body temperature) will be presented at study baseline, and prior and after injection by scheduled visit and treatment group.

[redacted]

Anti-nuclear antibody (ANA) status will be categorized as negative and positive that correspond

[redacted]

7.8 Additional Analyses to be Included in 24 Week CSR

In addition to safety summaries specified above for the 24-week double-blind period; summaries of all AEs, AEs of special interest (SAEs, Malignancies, Myositis related AEs and autoimmune events) reported in the database up to the 24-week database lock will be included in the 24-week study report. These AEs will be presented by the double-blind treatment groups for the cumulative abatacept period based on the as-treated analysis population.

8 CONVENTIONS

8.1 Calculations of Key Measures

8.1.1 IMACS DOI

IMACS DOI for the primary endpoint is defined as meeting all the following criteria:

- An improvement of ≥ 20% in 3 IMACS core measures, AND
- No more than 2 IMACS core measure scores worsen by ≥ 25%, AND
- Manual Muscle Test (MMT-8) may not decrease by ≥ 25%.

For the exploratory analyses two modified IMACS DOI definitions will be considered:

First alternative definition

- A1: An improvement of ≥ 25% in 3 IMACS core measures, AND
- A2: No more than 2 IMACS core measure scores worsen by ≥ 30%, AND
- A3: Manual Muscle Test (MMT-8) may not decrease by ≥ 30%.

Second alternative definition

- B1: An improvement of ≥ 30% in 3 IMACS core measures, AND
- B2: No more than 2 IMACS core measures scores worsen by ≥ 35%, AND
- B3: Manual Muscle Test (MMT-8) may not decrease by ≥ 35%.

For the exploratory analyses the modified IMACS DOI definition using a 25% (and 30%) cutoff for improvement in 3 core measures, and no more than 2 core measures worsening by 30% (and 35%) and MMT-8 not decrease by 30% (and 35%) will be provided.

The IMACS Core measures are:

- Physician Global Assessment of Disease Activity (PGA).
- Patient (Subject) Global Assessment of Disease Activity (SGA).
- Manual Muscle Test (MMT-8).
- Health Assessment Questionnaire-Disability Index (HAQ-DI).
- Muscle Enzyme levels.
- Myositis Disease Activity Assessment Tool (MDAAT) Extramuscular Global Activity VAS (V-8 in the CRF).

8.1.2 Physician Global Assessment of Disease Activity (PGA)

Physician Global Assessment of Disease Activity (PGA) is measured using a 10cm visual analogue scale that can be performed by the investigator or sub-investigator.

8.1.3 Patient Global Assessment of Disease Activity (SGA)

Patient (Subject) Global Assessment of Disease Activity (SGA) is measured using a 10cm visual analogue scale that is performed by the subject.

8.1.4 Manual Muscle Test (MMT-8)

This tool assesses muscle strength using a 0-10 point scale. A set of 7 designated muscles is tested bilaterally plus axial (neck flexor) testing, so that the potential MMT-8 score ranges from 0-150.

The designated muscles tested will be:

- Neck flexors.
- Deltoid middle.
- Biceps brachii.
- Gluteus maximus.
- Gluteus medius.
- Quadriceps.
- Wrist extensors.

- Ankle dorsiflexors.

8.1.5 HAQ Disability Index

Scoring conventions for the Standard Disability Index of HAQ will be followed. The Standard disability index (HAQ DI) takes into account the subject's use of aids or devices or assistance in the scoring algorithm for a disability category. For each of the eight disability categories there is an AIDS OR DEVICES companion variable(s) that is used to record the type of assistance, if any, a subject uses for his/her usual activities. If aids or devices and/or assistance from another person are checked for a disability category, the score for this category is set to 2 (much difficulty), if the original score is 0 (no difficulty) or 1 (some difficulty). The HAQ DI is then calculated by summing the adjusted categories scores and dividing by the number of categories answered. Details of the HAQ DI scoring conventions are documented in "The Health Assessment Questionnaire".

8.1.6 Muscle Enzyme Levels

The muscle enzymes of interest are creatine kinase (CK), aldolase, AST, ALT, and LDH. To assess the DOI the most abnormal lab value (% above ULN) at baseline will be determined. The subsequent change from baseline for this lab value will be used to determine if the patient has met the criteria for the IMACS Definition of Improvement (DOI) at each time point. To assess the criteria for worsening, the change from baseline point will be determined for all 5 lab values at each time point indicated. The lab value with the greatest percent increase from baseline for each will be used to determine if the patient has met the criteria for worsening respectively (Section 3.1.5 of the Protocol).

8.1.7 Extramuscular Disease Activity using Myositis Disease Activity Assessment Tool (MDAAT)

This is a combined tool that captures the physician's assessment of disease activity of various organ systems using (1) a 0-4 scale and (2) a visual analog scale (VAS). It assesses the clinical features of each organ system based upon:

- The presence of clinical features or symptoms within the previous 4 weeks that are due to active disease (i.e., use your clinical judgment to determine how active the myositis associated clinical feature has been within the previous 4 weeks).
- The judgment that the feature is due to the myositis disease process (i.e., clinical findings known or suspected to be due to another disease process or due to therapy should NOT be considered in this evaluation).
- The concept that disease activity is defined as a potentially reversible finding.
- A clinical, functional, and laboratory assessment for each organ system.

The scoring system is based primarily on the physician's intention to treat (Categories A – D). In some cases, a patient may meet the clinical manifestations of a particularly category, but the treatment does not match the intention to treat of that particular category; in those cases, the category matching the clinical symptoms should be marked. If more than one clinical symptom is

present within a system, and different categories would be scored for each, mark the category that is most severe.

8.1.8 IMACS Calculation

IMACS DOI will be evaluable if there are at least 5 non-missing core measures, including MMT-8. In that case IMACS DOI status will be categorized as a response based on the available set of core measures satisfying the criteria mentioned above, otherwise, the status will be set as a non-response. If MMT-8 is not available or less than 5 core measures are available then IMACS DOI status will be subject to the imputation methods specified in Section 8.1.9.

Higher values for each component of the IMACS core measures (excluding MMT-8) indicate greater severity of disease. Improvement from baseline in a core measure of the IMACS will be computed as the difference between the baseline (Day 1 pre-treatment) value and the value at a given post-baseline visit. A positive value for improvement from baseline for an individual core measure indicates lesser severity of disease. The IMACS DOI will be defined based on the percent improvement from baseline in each of the IMACS core measure listed in Section 8.1.1.

Percent improvement will be computed as the ratio of improvement from baseline to the baseline value times 100 for an endpoint:

$$\% \text{ Improvement} = 100 * (\text{Improvement from Baseline}) / \text{Baseline Value}$$

$$= 100 * (\text{Baseline Value} - \text{Post-baseline Value}) / \text{Baseline Value}$$

For MMT-8, a higher value implies a lesser severity of disease; the % reduction in muscle strength will also be calculated as:

$$100 * (\text{Baseline Value} - \text{Post-baseline Value}) / \text{Baseline Value}$$

8.1.9 Determination of IMACS DOI Responders

Any subject who does not meet the IMACS DOI criteria is considered a non-responder. For IMACS DOI status that cannot be evaluated due to missing data (i.e. MMT-8 not available or less than 5 of the core measures available) the following conventions will be implemented:

- For the primary analysis at Week 24, any subject who prematurely discontinues the trial after receiving study medication will have missing data imputed as a IMACS/DOI non-responder at all scheduled protocol visits subsequent to the point of discontinuation. For the sensitivity analysis of the primary endpoint the data collection post discontinuation from the study will be used in the determination of the DOI status and included in the analyses. The tipping point sensitivity analysis will impute the missing data as described in Section 7.5.3.2
- Any subject for whom data are missing at a given visit not due to premature discontinuation will have a value for IMACS/DOI response imputed at that given visit. If the given visit is not the last scheduled efficacy visit (Day 169), imputation will depend on the observed responses from the previous scheduled visit and the next scheduled visit. If there are IMACS DOI responses observed at both visits, an IMACS DOI response will be imputed for the given

missing visit. Else, if a non-response is observed at one of the visits or a response is missing at one of the visits then a non-response will be imputed. If the given visit is the last scheduled efficacy visit (Day 169), the missing response will be imputed as non-response.

8.1.10 Myositis Function Index (FI-2)

The FI-2 is a functional outcome developed for patients with adult polymyositis or dermatomyositis assessing muscle endurance in seven muscle groups. Each muscle group is scored as the number of correctly performed repetitions with 60 or 120 maximal number of repetitions depending on muscle group. The FI-2 is a further development of the original Functional Index (FI) and has been validated as to content and construct validity and intra- and inter-rater reliability. The FI-2 can be performed on both right and left sides requiring a maximum of 33 minutes, or just on the dominant side which takes 21 minutes. If both sides are available for a given muscle set then the scores from both sides will first be averaged so that there is only one score per muscle set for the calculations below. If only one side is available then that side will be used.

The 3 types of measures for analysis will be as follows:

- Total score based on hip flexion, shoulder flexion (R/L) and neck divided by 3.
- Total score based on all 7 muscle sets.
- Individual scores for all 7 muscle sets.

8.1.11 Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI)

This tool assesses the improvement in skin disease activity in subjects with dermatomyositis. Four domains are evaluated including 1) disease activity and damage in 15 anatomic locations, 2) presence of Gottron's papules on the hands, 3) periungual changes, and 4) alopecia. A total activity score and damage score are calculated.

8.1.12 Myositis Damage Index (MDI)

This tool measures the degree of disease damage of all organ systems. It is composed of a series of organ-specific questions relating to the presence or absence of a given sign or symptom or problem to measure the extent of damage, and an overall rating of the disease damage of each system using a 10 cm visual analogue scale to measure the severity of damage.

The MDI has 3 proposed scores: an extent of damage score, a severity of damage score and an extended damage score.

The proposed scoring system for the MDI extent of damage score is the sum of all 0 or 1 scores for the 11 individual organ systems (muscle, skeletal, cutaneous, gastrointestinal, pulmonary, cardiovascular, peripheral vascular, endocrine, ocular, infection, malignancy) divided by the total possible score (range = 0–35 in children, 0–37 for adolescents and 0–38 in adults). If one or more items were not assessed, the resulting score would be divided by the maximum possible score of

the assessed items. The categories of "other damage" and "global damage" are not included in the MDI extent of damage score but are scored separately.

The proposed MDI severity of damage score is the sum of the 10 cm visual analogue scale scores for each of the 11 individual organ systems divided by the total possible score (range = 0–110). If one or more organ systems were not assessed, the resulting score would be divided by the maximum possible score of the assessed items. The categories of "other damage" and "global damage" are not included in the MDI severity score but are scored separately.

The proposed MDI extended damage score is the sum of the optional items listed in italics under the systems muscle, skeletal, gastrointestinal, pulmonary, endocrine divided by the total possible score. Each optional item is scored 0 or 1, providing a range of 0–16 for the extended damage score. If one or more items were not assessed, the resulting score would be divided by the maximum possible score assessed.

8.1.13 Myositis Response Criteria

The Myositis Response Criteria (MRC) is a novel outcome measure derived by consensus for a conjoint analysis–based continuous model using absolute percent change in the IMACS DOI core set measures. A total improvement score (range 0–100) can be determined by summing scores for each core set measure based on improvement adjusted by relative weight for each core set measure (see APPENDIX 2). A total improvement score of ≥ 20 represents minimal improvement, a score of ≥ 40 represents moderate improvement, and a score of ≥ 60 represents major improvement.

8.2 Baseline Measures

For all outcome measures, the baseline value is the last assessment taken prior to first dose of study medication. In general, this is the assessment taken on study Day 1 before the administration of study medication. For some variables that are also measured during screening, if Day 1 assessments are unavailable then the last assessments prior to Day 1 from screening assessments will be used as baseline value. When there is a missing baseline assessment in joint counts, it will not be imputed, thus, subjects are excluded from any change from baseline analysis for which they have a missing baseline value. No screening assessment will be imputed as baseline measurements for subjects who require a washout before the start of the study medication.

8.3 Missing Measurements

For listings of IMACS DOI response, imputed DOI response will be included in the listings with a flag to indicate the values were imputed; missing IMACS components will be presented as missing. For determination of IMACS DOI response, handling of missing values is addressed in Section 8.1.9.

For the analysis of individual components of IMACS DOI, HAQ DI and its category scores, FI-2, Myositis Response Criteria missing values will not be imputed. Subjects with only baseline

observations will be excluded from any changes from baseline analyses. Similarly for all other efficacy, safety, and immunogenicity measures and biomarkers, missing values will not be imputed.

8.4 Missing, Unknown or Partial Dates

The BMS safety guidelines for conventions relating to the handling of missing or partial dates and the determination of appropriate default values in such cases (in particular, for concomitant medication dose start-dates and end-dates and AE onset dates) will be utilized.

8.5 Day Ranges for Analysis of Time Points

Subjects do not always adhere strictly to the visit schedule timing in the protocol. Therefore the designation of visits during the double-blind period of the study will be based on the day of evaluation relative to the trial (day of first study medication = study Day 1) rather than the nominal visit recorded in the CRF. Mutually exclusive relative day windows are defined below to provide derived visits that correspond to the post-baseline time points. If a visit falls outside of the pre-specified visit windows, then the data collected at that visit will not be assigned a derived visit but will remain in the derived data sets. Determination of baseline values is addressed in Section 8.2.

Designation of visits for efficacy or laboratory assessments performed at every scheduled visit is tabulated in Table 8.5-1.

Table 8.5-1: Day Ranges for Efficacy or Laboratory Assessments During the Double-Blind Period

Visit	Target Day	Assessments
Screening	< 1	< 1
Day 1 (Baseline)	1	1**
Day 29	29	8 - 43
Day 57	57	44 - 71
Day 85 (Week 12)	85	72 - 99
Day 113	113	100 - 127
Day 141	141	128 - 155
Day 169 (Week 24)*	169	156 - 211

*Note: For subjects who continue into the open-label period of study, the upper limit of this window is the first dose day of the open-label period.

**Note: See Section 8.2 for baseline measures.

Designation of visits for the ANA assessments is tabulated in Table 8.5-2.

Table 8.5-2: Day Ranges for ANA Assessments at Every Scheduled Visit During the Double-Blind Period

Visit	Target Day	Assessments
Day 1 (Baseline)	1	≤ 1
Day 169 (Month 6)*	169	8 - 211

*Note: For subjects who continue into the open-label period of study, the upper limit of this window is the first dose day of the open-label period.

Designation of visits for the immunogenicity assessment by treatment group is tabulated in Table [redacted]

[redacted]

For subjects who discontinue from study therapy prematurely, the efficacy assessments performed after the last dose days of study drug will be included in the efficacy data sets provided that these assessments are made within 42 days of the last dose. Exceptions to this rule apply in specified sensitivity analyses that include off-treatment assessments for subjects who discontinued treatment early.

If a subject has more than 1 visit where a measurement is recorded within a window, the measurement closest to the target day will be used. In case of 2 visits being equidistant from the target, the later measurement will be used in analyses. Exceptions to these rules apply to ANA and immunogenicity. For these safety indicators, the least favorable value (toward a positive response) in the window will be used.

9 CONTENT OF REPORTS

The results of this study will be presented in a standard BMS Clinical Study Report (CSR). Prior to completion of the CSR and Initial Data Assessment will be prepared briefly identifying the key results and any unanticipated findings that are unusual for a study within this program.

10 DOCUMENT HISTORY

Table 10-1: Document History

Version Number	Author(s)	Description
1.0	[redacted]	Original Issue

APPENDIX 1 RELEVANT PROTOCOL DEVIATIONS

Pre-randomization

- Myositis Activity:
  - Screening MMT-8 score of 140 or more.
  - Failure to document criteria for currently active IIM or obtain approval from adjudication committee.
  - Diagnosis of Inclusion Body Myositis.
- Medication/Procedure/Other:
  - Use of rituximab within 3 months.
  - Use of immune globulin within 6 weeks.
  - Discontinuation of prohibited medication within 2 weeks.

Post-randomization

- Medication/Procedure:
  - Use at randomization of more than 1 of the allowable immunosuppressant's drugs in the Double-Blind period.
  - Use of corticosteroids above prednisone ≥ 35mg/day (or prednisone equivalent) for > 3 weeks or within 2 weeks of Day 169.
  - Increase of 5 mg or more per day or initiation of more than 10 mg per day of prednisone (or prednisone equivalent) that is unrelated to management of an AE or permitted rescue therapy.
  - All prohibited medication not discontinued at least 2 weeks prior to randomization with the exception of immune globulin (3 months), rituximab (6 months) and d-penicillamine (3 months). Prohibited medication in this context refers to immunomodulatory medications for the treatment of IIM which are not included as immunosuppressant medications. The following list includes examples of specific prohibited medications:
    - Adrenal Corticotropic Hormone (ACTH), including analogues.
    - Sirolimus.
    - Everolimus.
    - Mizoribine.
    - Tofacitinib or other JAK inhibitor.
  - All investigational or approved biologic DMARD therapies other than abatacept (including but not limited to tocilizumab, adalimumab, golimumab, certolizumab, etanercept, infliximab, anakinra, rituximab, etc.).
  - Use of any investigational drug other than study medication.
- Study conduct:
  - Subject missed more than 6 consecutive doses of study medication (abatacept or placebo) during the Double-Blind period or missed more than 4 consecutive scheduled doses of study medication (abatacept or placebo) between Day 141 and Day 169.

Covid-19 Related Deviations

- Unable to attend both Day 141 and Day 169 assessments due to travel/other restrictions.
- For missed doses due to Covid-19 the same rule to define those as a relevant protocol deviation applies as under study conduct above.

APPENDIX 2 MYOSITIS RESPONSE CRITERIA

The final myositis response criteria for minimal, moderate and major improvement in adult dermatomyositis/ polymyositis (DM/PM) and combined DM/PM and juvenile DM clinical trials and studies.

Table 10-2: Myositis Response Criteria

Core set measure, level of improvement based on absolute percent change	Improvement score
Physician global activity	
Worsening to 5% improvement	0
>5% to 15% improvement	7.5
>15% to 25% improvement	15
>25% to 40% improvement	17.5
>40% improvement	20
Patient global activity	
Worsening to 5% improvement	0
>5% to 15% improvement	2.5
>15% to 25% improvement	5
>25% to 40% improvement	7.5
>40% improvement	10
Manual muscle testing	
Worsening to 2% improvement	0
>2% to 10% improvement	10
>10% to 20% improvement	20
>20% to 30% improvement	27.5
>30% improvement	32.5
Health Assessment Questionnaire	
Worsening to 5% improvement	0
>5% to 15% improvement	5
>15% to 25% improvement	7.5
>25% to 40% improvement	7.5
>40% improvement	10
Enzyme (most abnormal)	
Worsening to 5% improvement	0
>5% to 15% improvement	2.5
>15% to 25% improvement	5
>25% to 40% improvement	7.5
>40% improvement	7.5
Extramuscular activity	
Worsening to 5% improvement	0
>5% to 15% improvement	7.5
>15% to 25% improvement	12.5
>25% to 40% improvement	15
>40% improvement	20

The total improvement score is the sum of all 6 improvement scores associated with the change in each core set measure. A total improvement score of ≥ 20 represents minimal improvement, a score of ≥ 40 represents moderate improvement, and a score of ≥ 60 represents major improvement. If one of the core measures is missing then the total score will be missing.

Improvement Score Calculation

The absolute percent change ((post-baseline value – baseline value) / range) * 100 is calculated for each core set measure. For muscle enzymes, the most abnormal serum muscle enzyme level at baseline (creatine kinase, aldolase, alanine transaminase, aspartate aminotransferase, lactate dehydrogenase) is used.

The range for each of the core set measures are as follows:

Muscle Enzymes: The enzyme range was calculated based on a 90% range of enzymes from natural history data as follows:

Creatinine Kinase: Range = 15 times the upper limit of normal (ULN)

Aldolase: Range = 6 times the ULN

Lactate dehydrogenase, aspartate aminotransferase, and alanine transaminase: Range = 3 times the ULN.

The ULN is based on what is reported in the lab dataset for each enzyme.

Ranges for other core set measures:

Physician Global Assessment of Disease Activity (PGA) VAS: Range = 10

Patient Global Assessment of Disease Activity (SGA) VAS: Range = 10

Manual Muscle Testing (MMT-8): Range = 150

HAQ: Range = 3

Extramuscular Global Activity (using MDAAT): Range = 10

An improvement score is assigned for each core set measure based on the absolute percent change in the core set measure according to the definition in Table 2.1 above. These individual core set measure improvement scores are then totaled among the 6 core set measures to give the total improvement score.

A total improvement score between 0 and 100 corresponds to the degree of improvement, with higher scores corresponding to a greater degree of improvement.

APPENDIX 3 ANALYSIS PLAN FOR OPEN-LABEL, OPEN-LABEL EXTENSION AND LONG-TERM EXTENSION PERIODS ANALYSIS

Analyses specified in this section will be performed:

1) At the end of the open-label period, i.e. after all subjects have completed the 52-week assessment.
2) At the end of the open-label extension period, i.e. after all subjects have completed the 76-week assessment (Japan sites only).
3) At the end of the long-term extension period, i.e. after all subjects have completed the 3-year assessment and all protocol specified follow-up visits.

For sections "Background and Rationale; Study Description; Objectives; Endpoints; Sample Size and Power; Study Periods, Treatment Regimens and Populations for Analyses; Conventions" and their subsections, see Section 1-Section 6 and Section 8 of SAP for 24-week analysis of the double-blind study period. Any changes or modification to the analysis for the double-blind period will be explicitly presented.

The open-label analysis will include safety assessment for all safety data in the database at the time of the 52-week database lock.

The open-label extension analysis will include safety assessment for all safety data in the database at the time of the 76-week database lock.

The long-term extension analysis will include safety assessment for all safety data in the database at the time of the study final database lock.

11 STATISTICAL ANALYSES

11.1 General Methods

All subjects who were randomized and treated during the open-label period will be included in the open-label analysis. The analyses will be based on the open-label treated analysis population and presented by the double-blind treatment group except otherwise stated.

All Japanese subjects who were randomized and treated during the open-label extension period will be included in the open-label extension analysis. The analyses will be based on the Japan open-label extension treated analysis population and presented by the double-blind treatment group except otherwise stated.

All subjects who were randomized and treated during the long-term extension period will be included in the long-term extension analysis. The analyses will be based on the long-term extension treated analysis population and presented by the double-blind treatment group except otherwise stated.

[redacted]

[redacted]

11.3 Study Conduct

The relevant protocol deviations are only assessed for the double-blind period and not for the other periods.

11.4 Study Population

Frequency distributions or summary statistics of data pertaining to subject disposition, demographics, baseline characteristics, and medical history will be tabulated and displayed by randomized treatment group for all subjects in the open-label / open-label extension / long-term extension periods. No statistical test will be carried out for comparison of any baseline measurement among treatment groups.

11.4.1 Subject Disposition

The number of subjects randomized, the number of subjects treated in the open-label / open-label extension / long-term extension periods, and the number of subjects who discontinued the open-label / open-label extension / long-term extension periods, together with reasons for discontinuation taken from the eCRF status page, will be tabulated by randomized treatment group.

11.4.2 Demography and Baseline Characteristics

Demographic and baseline characteristics of all subjects in the open-label / Japan open-label extension / long-term extension treated analysis populations; summaries will be presented by the double-blind treatment groups. Continuous variables such as age, weight, and duration of disease will be summarized by randomized treatment group using means, standard deviations and ranges. The distribution of categorical variables such as gender, race and geographic region will be summarized by treatment group using frequencies and percentages. See Section 7.3.2 for variables included in these summaries. This will be based on the data captured on the CRF page, except for the summary of the stratification variables which is based on the IxRS data.

11.4.3 Medical History and Prior Medication

Medical history and prior medication information before the first study medication will be summarized by treatment group using relative frequencies in open-label / Japan open-label extension / long-term extension treated analysis populations. Listings of medical history and prior medication will be presented by treatment group in the open-label / Japan open-label extension / long-term extension treated analysis populations.

11.5 Extent of Exposure

11.5.1 Study Therapy

Extent of exposure to study drug in the open-label / Japan open-label extension / long-term extension treated analysis populations will be summarized in two ways; first, by the number of abatacept injections, summarizing frequency distribution of number of abatacept injections taken by treatment group, and second, by the number of days (or months) the subject is known to be on abatacept, ignoring any dosing interruptions. Exposure will be presented separately for the open-label / open-label extension / long-term extension periods and the cumulative abatacept period.

11.5.1.1 Open-Label Period Exposure

The extent of exposure for the open-label period will be calculated and presented in days as follows:

- For subjects who discontinued during the open-label period or did not enter the open-label extension period or the long-term extension period within 56 days of the last injection in the open-label period:
- Exposure in days = (date of last injection in the open-label period - date of first open-label injection + 1) + 56.

For subjects who enter open-label extension or long-term extension period within 56 days of the last injection in the open-label period:

- Exposure in days = (date of first injection in the open-label extension or long-term extension period - date of first open-label injection).

The offset of 56 days is the length of 2 regular dosing cycles and represents approximately 4 times the half-life of Abatacept in humans.

Summaries of exposure to study drug will show the distribution of the number of injections and days on drug, together with the means, standard deviations, medians, minima and maxima, by treatment group.

11.5.1.2 Open-Label Extension Period Exposure

The extent of exposure for the open-label extension period will be calculated and presented in days as follows:

- For subjects who discontinued during the open-label extension period or did not enter the long-term extension period within 56 days of the last injection in the open-label extension period:
- Exposure in days = (date of last injection in the open-label extension period - date of first open-label extension injection + 1) + 56.

For subjects who enter long-term extension period within 56 days of the last injection in the open-label extension period:

- Exposure in days = (date of first injection in the long-term extension period - date of first open-label extension injection).

11.5.1.3 Long-Term Extension Period Exposure

The extent exposure of for the long-term extension period will be calculated and presented in days as follows:

Exposure in days = (date of last injection in the long-term extension period - date of first long-term extension injection + 56).

11.5.1.4 Cumulative Abatacept Exposure

The cumulative abatacept exposure will only be computed for the final analysis (i.e. at the study closeout); the exposure to abatacept during the cumulative period will be calculated as follows:

Exposure in months = (date of last dose of abatacept in the study - date of first dose of abatacept in the study + 1 - adjustment + 56)/30.

The offset of 56 days represents approximately 4 times the half-life of abatacept in humans.

Adjustment is the sum of the time span between:

- the last dose of abatacept in the double-blind period and the first dose of abatacept in the open-label period which exceeds 56 days (all sites),
- the last dose of abatacept in the open-label period and the first dose of abatacept in the open-label extension period which exceeds 56 days (Japan sites); or the last dose of abatacept in the open-label period and the first dose of abatacept in the long-term extension period which exceeds 56 days (all other sites but not including U.S. and Czech Republic sites),
- the last dose of abatacept in the open-label extension period and the first dose of abatacept in the long-term extension period which exceeds 56 days (Japan sites).

Summaries of cumulative period exposure to study drug will show the distribution of the number of months on drug, together with the means, standard deviations, medians, minima and maxima, by treatment group. A listing of abatacept study medication dose dates will also be provided for the cumulative period.

11.5.2 Discontinuation of Study Therapy

Discontinuation from study therapy is defined as a subject's termination of the study medication without resumption prior to study completion. The discontinuation from study therapy during the open-label / open-label extension / long-term extension periods will be summarized for all open-

label / open-label extension / long-term extension periods treated subjects, by treatment group and the reason for premature termination, taken from the subject status eCRF page.

11.5.3 Treatment Compliance

The number of subjects who missed study medication doses (excluding missed doses due to premature discontinuation from the study) during open-label / open-label extension / long-term extension will be summarized for all open-label / open-label extension / long-term extension periods treated subjects, by double-blind treatment group and number of study dose missed. A corresponding listing of all treated subjects with missed doses during the open-label / open-label extension / long-term extension periods will be provided.

[redacted]

11.6 Efficacy

All efficacy analyses for the open-label / open-label extension periods will be performed using the open-label / Japan open-label extension treated analysis population, unless otherwise specified. Over time (double-blind and open-label / open-label extension visits) analyses will be performed for all efficacy endpoints specified in the 24 week Analysis Plan. The analyses will be as observed with no imputation for any missing data. No p-values will be provided for any of the open-label analyses.

11.6.1 Other Efficacy Analyses

IMACS DOI

Proportion of subjects in IMACS DOI over time during the study will be presented by the double-blind treatment groups and visit. Point estimate of response rate, 95% CI, point estimates and 95% CI of treatment difference adjusted for randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs present), Region (Japan vs ROW) (based on minimum risk weights).

Proportion of subjects who meet the ≥ 20% improvement from baseline over time for each of the core measures of IMACS (PGA, SGA, MMT-8, HAQ-DI, MDAAT and muscle enzymes) will be presented by double-blind treatment groups and visit. Point estimate of response rate, 95% CI, point estimates and 95% CI of treatment difference adjusted for randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs present), Region (Japan vs ROW) (based on minimum risk weights).

The proportion of subjects in IMACS Minimal, Moderate and Major responses using the Myositis Response Criteria tool (as defined in section 8.1.11) will also be presented over time during the study. Point estimates and 95% CI within each treatment group for IMACS Minimal, Moderate and Major DOI response will be provided by visit. Differences in IMACS Minimal, Moderate and Major DOI response rates between the double-blind treatment groups with be provided by point estimate and corresponding 95% CIs based on an ordinal logistic regression model.

Additional separate longitudinal repeated mixed effects models will be used to assess the mean change from baseline over time in the study in each of the IMACS core measures (PGA, SGA, MMT-8, HAQ-DI, MDAAT and muscle enzymes). The mixed models will include double-blind treatment group, baseline measurement, randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs present), Region (Japan vs ROW), time (study days in which each core measures was collected), time by baseline measurement and time by treatment interactions as fixed effects and subject as a random effect. Adjusted mean, SE, 95% CI for adjusted mean difference between treatment groups will also be provided. The estimate of baseline mean core measure, mean post-baseline core measure, and adjusted mean change from baseline core measure will be also be presented by double-blind treatment group and visit.

11.6.2 Additional Efficacy and Outcome Research Analyses

FI-2

Changes in muscle in endurance measured by FI-2 scores over time in the study will be assessed using a longitudinal repeated measures mixed model. The model will include baseline FI-2 score, double-blind treatment group, randomization stratification factors disease diagnosis (DM vs. Non-DM), ILD (absent vs present), Region (Japan vs ROW), time and time by treatment interaction. The estimate of baseline mean FI-2 score, mean post-baseline FI-2 score, and adjusted mean change from baseline FI-2 score will be summarized by treatment group and visit. Treatment difference in adjusted mean changes from baseline, along with the corresponding 95% confidence intervals, will also be presented.

Myositis Response Criteria

To assess the myositis response criteria score over time during the study a longitudinal repeated mixed effects model will also be used. The mixed model will include double-blind treatment group and randomization stratification factors: disease diagnosis (DM vs. Non-DM), ILD (absent vs present), Region (Japan vs ROW), time (study days in which MRC is measured) and time by treatment interactions as fixed effects and subject as a random effect. Adjusted mean difference, SE, 95% CI for adjusted mean difference between treatment groups will be provided. The estimate of mean post-baseline MRC total improvement score will be also summarized by double-blind treatment group and visit.

[redacted]

post-baseline MDI, and adjusted mean change from baseline MDI score will be also be presented

[redacted]

11.7 Safety

Unless otherwise specified, the same safety analyses for the Week 24 analysis will be performed for AEs that occurred during the open-label / open-label extension / long-term extension periods.

Additional analyses will also be provided for the cumulative Abatacept period in subjects who received at least one dose of Abatacept at any time during the study. The analyses in addition to the safety analyses for Week 24 analysis described in Section 7.6 are explicitly summarized in the following:

- Incidence rates by 3-month intervals for SAEs, pre-specified infections and adverse events of interest (*Infections, Malignancy, Autoimmune disorders)* will be provided for cumulative abatacept period.
- Multiple events will be reported based on unique AEs for each specific preferred term. For each subject and preferred term, a single record (unique AE) is used to represent multiple AEs when:
  - Multiple AE records have the same onset date.
  - The onset date of an event record is either the same day or 1 day later than the resolution date of a preceding event record (continuous events).
  - The onset date of an event record is after the onset date and prior to or on the resolution date of a preceding event record (overlapping events).

The unique AE record contains the earliest onset date, latest resolution date (if available), highest intensity, latest relationship, the most severe adverse event type (i.e., the SAE if ever reported as SAE), treatment required (will be yes if ever required) and highest action taken in the following order (highest to lowest): drug discontinued, drug interrupted, dose reduced, dose increase, none.

Multiple events will be presented for the cumulative abatacept period based on subjects who received at least one dose of abatacept at any time during the study.

Exposure adjusted (per 100 person-years) incidence rates of adverse events accounting for multiple occurrence of an event will be calculated for all AEs/SAEs which occurred during the open-label / open-label extension / long-term extension periods in abatacept treated subjects. The numerator is the number of unique events within the open-label / open-label extension / long-term extension periods (up to 56 days post last abatacept dose during the open-label / open-label extension / long-term extension periods). The denominator is the overall total exposure (person-years) within each period, which is calculated as the sum over all subjects of exposure (in days) divided by 365.25. The resulting incidence rate is multiplied by 100 to express the rate per 100 person-years.

As an example, if 10 patients report 7 unique episodes of headache and had a combined cumulative exposure of 20 years to study medication, the incidence rate is reported as 7 / 20 * (100) or 35 cases per 100 patient years of exposure.

For all SAEs, exposure adjusted per 100 person-years rates will be summarized by time-intervals in months: 0-3 months, 3-6 months, 7-9 months, 10-12 months and > 12 months (time-intervals defined as Days 1-90, Days 91-180, Days 181-270, 271-360 and > 360 Days), respectively. In addition, the number of subjects who experienced a SAE, once or multiple times within this period will be presented. The above analyses will be performed for infection and local injection site reaction, respectively. A subject listing of unique AEs will also be provided.

Unique AEs and exposure adjusted multiple events summaries will also be provided separately for the cumulative period. The exposure adjusted per 100 person-years rates will be provided for SAEs, Infections, local injection site reaction. and will be summarized by 3 month time interval for the cumulative period: 0-3 months, 3-6 months, 7-9 months, 10-12 months and > 12 months (time-intervals defined as Days 1-90, Days 91-180, Days 181-270, 271-360 and > 360 Days), respectively.

[redacted]

[redacted]