Official Title of Study:

A Multi-Center, Randomized, Double-Blind, Placebo- and Active Comparator-Controlled Phase 3 Study to Evaluate the Efficacy and Safety of BMS-986165 in Subjects with Moderate-to-Severe Plaque Psoriasis

NCT Number: NCT03624127

Document Date (Date in which document was last revised): December 17, 2019

Statistical Analysis Plan

Sponsor:	Bristol-Myers Squibb
Protocol No:	IM011046
[redacted]	
Version Date:	17-Dec-2019
Version No.:	6.0

Title:	A Multi-Center, Randomized, Double-Blind, Placebo- and Active Comparator-controlled Phase 3 Study to Evaluate the Efficacy and Safety of BMS-986165 in Subjects with Moderate-to-Severe Plaque Psoriasis
[redacted]	
SAP No.	2.0

Approved v2.0 [redacted]

Approvals

[redacted] Author	
[redacted] Title:	[redacted]
Signature /Date:	

[redacted] Approvals	
[redacted] Title:	
Signature /Date:	

BMS Approvals	
[redacted]	[redacted]
Signature /Date	
[redacted]	
Signature /Date:	

Table of Contents

Abbreviations

Glossary of Abbreviations:	
AE	Adverse event
AEI	Adverse event of interest
ANCOVA	Analysis of covariance
[redacted]	[redacted]
ATC	Anatomic Therapeutic Classification
BID	Twice daily
BSA	Body surface area
CI	Confidence interval
CK	Creatine kinase
CMH	Cochran-Mantel-Haenszel
CRF	Case Report Form
CSP	Clinical Safety Program
CSR	Clinical Study Report
CTCAE	Controlled Terminology Criteria for Adverse Events
DLQI	Dermatology Life Quality Index
DMC	Data Monitoring Committee
EAIR	Exposure-adjusted incidence rate
ECG	Electrocardiogram
eC-SSRS	Electronic Columbia-Suicide Severity Rating Scale
[redacted]	[redacted]
FAS	Full analysis set
[redacted]	[redacted]
IL	Interleukin
IRS	Independent Reporting Statistician
IRT	Interactive Response Technology
ITT	Intention-to-treat
LOCF	Last observation carried forward
LS	Least-squares
MAR	Missing at random
MedDRA	Medical Dictionary for Regulatory Activities
mBOCF	Modified baseline observation carried forward
MI	Multiple imputation

Glossary of Abbreviations:	
[redacted]	[redacted]
[redacted]	[redacted]
NRI	Nonresponder imputation
PASE	Psoriatic arthritis screening and evaluation
PASI	Psoriasis Area and Severity Index
[redacted]	[redacted]
PDGD	Protocol deviation guidance document
PGA-F	Physician Global Assessment-Fingernails
[redacted]	[redacted]
[redacted]	[redacted]
PHQ-8	Eight-Item Patient Health Questionnaire
[redacted]	[redacted]
PP	Per-protocol
[redacted]	[redacted]
pp-PGA	Palmoplantar Physician's Global Assessment
PPS	Per-protocol set
PSSD	Psoriasis Symptoms and Signs Diary
[redacted]	[redacted]
QD	Once daily
QoL	Quality of Life
QTcB	QT interval corrected using Bazett's formula
QTcF	QT interval corrected using Fridericia's formula
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
[redacted]	[redacted]
sPGA	static Physician Global Assessment
ss-PGA	Scalp specific Physician's Global Assessment
TEAE	Treatment-emergent adverse event
TFLs	Tables, figures, and listings
[redacted]	[redacted]
[redacted]	[redacted]

1.0 Purpose

The statistical analysis plan (SAP) describes the statistical methods to be used during the reporting and analyses of data collected under BMS Protocol IM011046.

The SAP outlines the following:

- Study design
- Study objectives
- Endpoints and assessments
- Analysis sets
- Statistical methodology
- Conventions and definitions

The SAP should be read in conjunction with the study protocol and case report form (CRF) according to the version on Page 1 of this document. Any further changes to the protocol or CRF may necessitate updates to the SAP. Changes following approval of the first version of the SAP will be tracked in the SAP Change Log and a final version of the updated SAP will be approved prior to final database lock.

2.0 Study Description

2.1 Study Design

This is a 52-week, multi-center, randomized, double-blind, double-dummy, placebo- and active comparator-controlled study to evaluate the safety and efficacy of BMS-986165 vs placebo and apremilast. A total of 600 qualified subjects with moderate-to-severe plaque psoriasis will be enrolled.

Day 1 activities
Following a screening period of up to 4 weeks, qualified subjects who have completed the screening procedures and have met the inclusion/exclusion criteria will be randomized in a blinded manner in a 2:1:1 ratio via interactive response technology (IRT) to one of the following 3 treatment groups:

- BMS-986165 6 mg once daily (QD)
- Placebo
- Apremilast as an active comparator that is marketed in various countries. It will be titrated to 30 mg twice daily (BID) as follows:
    - Day 1: 10 mg tablet in the morning
    - Day 2: 10 mg tablet in the morning and evening
    - Day 3: 10 mg tablet in the morning and 20 mg tablet in the evening
    - Day 4: 20 mg tablet in the morning and the evening
    - Day 5: 20 mg tablet in the morning and 30 mg tablet in the evening
    - Day 6 and thereafter: 30 mg tablet in the morning and the evening

Dummy tablets (placebo for the BMS-986165 6 mg tablet, placebo for apremilast 30 mg tablet BID, and placebo for apremilast 10 mg, 20 mg, 30 mg during titration) will be administered to the subjects to maintain blinding in a double-dummy fashion. Note that apremilast will not be used as a treatment arm in China due to lack of market approval of the agent in China.

Week 16 activities
The coprimary endpoints, static Physician Global Assessment (sPGA) 0/1 and Psoriasis Area and Severity Index (PASI) 75 will be assessed at Week 16. Subjects receiving placebo will be switched in a blinded manner to BMS-986165 6 mg QD at Week 16. Subjects who are randomized to BMS-986165 6 mg QD or apremilast will continue on their assigned treatment regimen in a blinded manner.

Week 24 activities

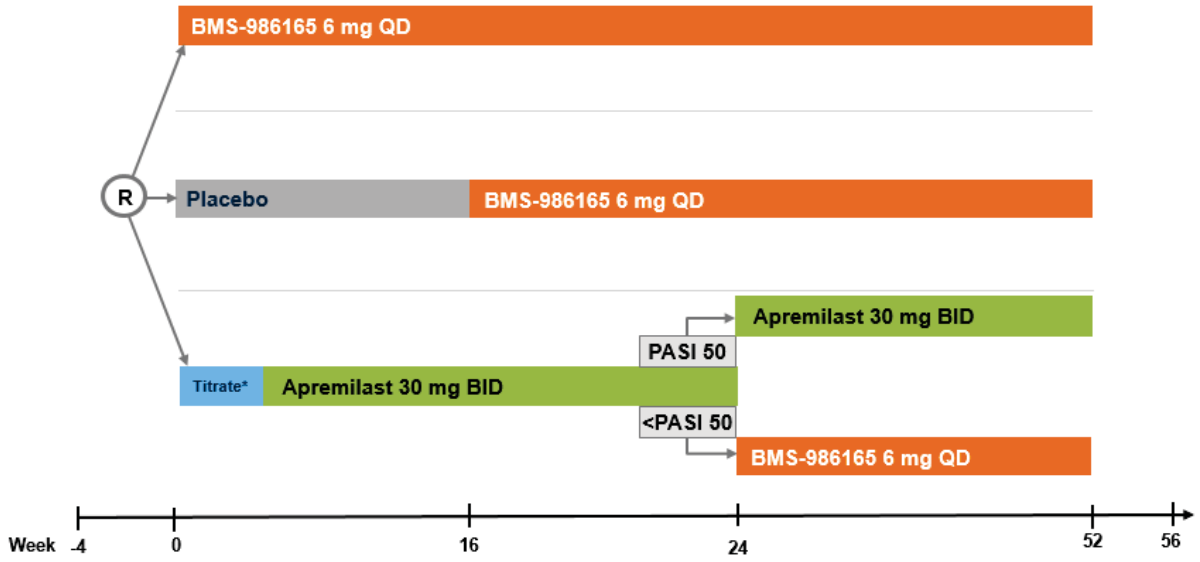
At Week 24, subjects originally randomized to apremilast who do not achieve PASI 50 response will be switched in a blinded manner to BMS-986165 6 mg QD. Subjects in the apremilast treatment arm who achieve PASI 50 response at Week 24 will continue to receive apremilast in a blinded manner through Week 52.

During the Week 24 assessment, a subject who has an sPGA ≥3 or ss-PGA≥3 may be treated with restricted topicals/shampoos as described in the protocol (Section 6.7.1). These treatments may only be initiated at Week 24, and not at subsequent time points. A subject who is provided these treatments at Week 24 may use them as needed per the investigator's judgement through Week 52.

Study Design elements

The duration of study participation is approximately 60 weeks and will be divided into the following periods: Screening (up to 4 weeks), Treatment (52 weeks), and Follow-up (4 weeks). A schedule of assessments can be found in the protocol. A study design schematic is provided in Figure 1.

Figure 1: Study Design Schematic



[redacted]

2.3 Treatment Assignment and Randomization

At Week 0 (Day 1), subjects who have met all criteria for enrollment will be centrally randomized by a computer-generated randomization schedule in a 2:1:1 ratio to the following treatments:

- BMS-986165 6 mg QD
- Placebo
- Apremilast 30 mg BID (titrated)

The randomization list was generated by the IRT vendor using a permuted block design within each stratum combination level. Randomization will be stratified by geographic region (U.S., Japan [body weight stratum will not be applied in Japan], China [body weight stratum will not be applied in China], Rest of World), previous biologic use (for psoriasis, psoriatic arthritis, or other inflammatory disease only; yes/no), and body weight (≥90 kg and <90 kg).

A treatment group will be assigned by IRT based on the above-described randomization schedule and each subject will be assigned a unique randomization number. In addition, a kit (container) number will be assigned to the subject by the IRT each time study treatment is dispensed. Dummy tablets (placebo to the BMS-986165 6 mg tablet and placebo to apremilast) will be administered to the subjects to maintain blinding. Apremilast will not be used as a treatment arm in China.

At Week 16, subjects receiving placebo will be switched in a blinded manner to BMS-986165 6 mg QD.

At Week 24, subjects originally randomized to apremilast who do not achieve a PASI 50 response will be switched in a blinded manner to BMS-986165 6 mg QD. Subjects originally randomized to apremilast who achieve PASI 50 response at Week 24 will continue to receive apremilast in a blinded manner. The investigative site and other study personnel will not have knowledge of the PASI 50 score at this visit and will therefore remain blinded.

2.4 Unblinding Information

The Data Monitoring Committee (DMC) provides oversight of safety considerations throughout the study. A separate unblinded team, comprised of an unblinded Independent Reporting Statistician (IRS) and unblinded programmer(s), will produce output for the DMC using masked treatments. Treatment decodes may only be requested by the DMC Chair and will be provided by the IRS. Data summaries and listings will be transmitted via a secure portal by the IRS to only the DMC members. Additional details regarding the DMC process and unblinding are provided in the DMC charter.

[redacted]

2.5 Changes in Statistical Considerations from the Protocol

The following is a list of the important changes in the SAP from the Statistical Considerations section in the protocol:

- The hierarchical testing order for the key secondary endpoints has been updated and two separate hierarchies have been provided, one for US submission and one for ex-US submission (see Tables 1 and 2 in Sec. 6.1.3). The update to the hierarchies are due to emerging information from clinical trials conducted with other agents recently approved for the treatment of psoriasis. The hierarchies presented in this document supersede the one that is in the protocol.

- Key secondary endpoints were added for the comparison of BMS-986165 to apremilast for sPGA 0/1, PASI 75, and PASI 90 at Week 24.
- Imputation methods were updated to remove the prohibited medication/therapy criteria for binary and continuous endpoints.
- The Full Analysis Set was changed from “all subjects who were randomized to receive assigned study treatment” in the protocol to “all subjects who are randomized”.
- The list of relevant protocol deviations provided in the protocol was updated for determination of the Per Protocol population.
- Logistic regression analyses for binary endpoints were removed as these are similar to the CMH analyses.

3.0 Objectives

3.1 Primary Objective

- Assess whether BMS-986165 is superior to placebo at Week 16 in the treatment of subjects with moderate-to-severe plaque psoriasis

3.2 Secondary [redacted] Objectives

- Assess whether BMS-986165 is superior to apremilast at Week 16
- Assess whether BMS-986165 is superior to apremilast at Week 52
- Assess whether BMS-986165 is superior to placebo over the first 16 weeks of treatment
- Assess whether BMS-986165 is superior to apremilast over 52 weeks of treatment
- Assess whether BMS-986165 is superior to placebo in scalp psoriasis through Week 16 in those subjects who have baseline scalp severity Physician’s Global Assessment (ss-PGA) score ≥3
- [redacted]
- Assess whether BMS-986165 is superior to placebo in nail psoriasis through Week 16 in those subjects who have baseline Physician’s Global Assessment-Fingernail (PGA-F) psoriasis score ≥3
- [redacted]
- Assess whether BMS-986165 is superior to placebo in palmoplantar psoriasis through Week 16 in those subjects who have baseline palmoplantar Physician’s Global Assessment (pp-PGA) score ≥3
- Evaluate improvement in patient-reported outcomes for BMS-986165 compared with placebo through Week 16
- Evaluate improvement in patient-reported outcomes for BMS-986165 compared with apremilast through Week 52
- [redacted]

4.0 Outcomes

The description of assessments for efficacy and safety can be found in Section 8 of the protocol. The calculation of key measures are provided in Section 8.2 of the SAP.

4.1 Efficacy

4.1.1 Primary Endpoint(s)

The coprimary endpoints for BMS-986165 compared to placebo at Week 16 are defined as:

- sPGA 0/1 response assessed as a proportion of subjects with an sPGA score of 0 or 1 with at least 2-point improvement from baseline.
- PASI 75 response assessed as a proportion of subjects who achieve a 75% improvement from baseline in the PASI score

4.1.2 Secondary Endpoint(s)

4.1.2.1 Key Secondary Endpoints for Comparisons to Placebo

The key secondary efficacy endpoints for BMS-986165 compared to placebo at Week 16 are defined as:

- PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in PASI score
- sPGA 0 response assessed as a proportion of subjects with an sPGA score of 0
- Scalp specific Physician's Global Assessment (ss-PGA) 0/1 assessed as a proportion of subjects with a ss-PGA score of 0 or 1 with at least a 2-point improvement from baseline among subjects with a baseline ss-PGA score ≥3
- PSSD symptom score of 0 assessed as a proportion of subjects with a PSSD symptom score of 0 among subjects with a baseline PSSD symptom score ≥1
- PASI 100 response assessed as a proportion of subjects who achieve a 100% improvement from baseline in PASI score
- Physician Global Assessment-Fingernails (PGA-F) 0/1 assessed as a proportion of subjects with a PGA-F score of 0 or 1 with at least a 2-point improvement from baseline among subjects with a baseline PGA-F score ≥3
- DLQI 0/1 assessed as a proportion of subjects with a DLQI score of 0 or 1 among subjects with a baseline DLQI score ≥2 (Ex-US submission only)

4.1.2.2 Key Secondary Endpoints for Comparisons to Apremilast

The key secondary endpoints for BMS-986165 compared to apremilast at Week 16 are defined as:

- sPGA 0/1 response assessed as a proportion of subjects with an sPGA score of 0 or 1 with at least a 2-point improvement from baseline
- PASI 75 response assessed as a proportion of subjects who achieve a 75% improvement from baseline in the PASI score
- PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in the PASI score
- Change from baseline in PSSD symptom score
- ss-PGA 0/1 assessed as a proportion of subjects with a ss-PGA score of 0 or 1 with at least a 2-point improvement from baseline among subjects with a baseline ss-PGA score ≥3
- sPGA 0 response assessed as a proportion of subjects with an sPGA score of 0
- PSSD symptom score of 0 assessed as a proportion of subjects with a PSSD symptom score of 0 among subjects with a baseline PSSD symptom score ≥1

The key secondary endpoints for BMS-986165 compared to apremilast at Week 24 are defined as:

- sPGA 0/1 response assessed as a proportion of subjects with an sPGA score of 0 or 1 with at least a 2-point improvement from baseline at Week 24
- PASI 75 response assessed as a proportion of subjects who achieve a 75% improvement from baseline in the PASI score at Week 24
- PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in the PASI score at Week 24

The key secondary endpoints for BMS-986165 compared to apremilast through Week 52 are defined as:

- sPGA 0/1 response assessed as a proportion of subjects with an sPGA score of 0 or 1 with at least a 2-point improvement from baseline at Week 52 and at Week 24
- PASI 75 response assessed as a proportion of subjects who achieve a 75% improvement from baseline in the PASI score at Week 52 and at Week 24
- PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in the PASI score at Week 52 and at Week 24

[redacted]

[redacted]

[redacted]

4.2 Safety

The safety outcomes include the following:

- Adverse events (AEs)
  - Treatment-emergent adverse events (TEAEs) – defined as:
    - AEs which occur after the first dose of study treatment through 30 days after the final dose of the study treatment or subject's participation in the study if the last scheduled visit occurs at a later time.
  - Treatment-emergent adverse events of interest (AEIs) for the following events:
    - Skin-related AEs
    - Infection AEs, including influenza
    - Creatine kinase (CK) elevation (evaluated as lab toxicity grade 2 or higher)
    - Malignancy
  - SAEs
  - Deaths
- Clinical laboratory parameters
  - Absolute and change from baseline values
  - Laboratory abnormalities (as determined by Controlled Terminology Criteria for Adverse Events [CTCAE v5.0] grading) presented as the worst postbaseline toxicity group
  - Shifts from baseline to maximum postbaseline value
  - Potential drug induced liver injury (DILI) is defined as a subject who meets the following criteria:
    - 1) ALT or AST elevation >3 times ULN

    AND

    - 2) Total bilirubin >2 times ULN, without initial findings of cholestasis (elevated serum alkaline phosphatase)

    AND

    - 3) No other immediately apparent possible causes of liver function test elevation and hyperbilirubinemia, including, but not limited to, viral hepatitis, preexisting chronic or acute liver disease, or the administration of other drug(s) known to be hepatotoxic.
  - ALT or AST elevation >5 times ULN
- Vital signs
  - Absolute and change from baseline values
  - Marked abnormalities defined by the below categories:
    - Heart rate:
      - Value > 100 and change from baseline > 30
      - Value < 55 and change from baseline < -15
    - Systolic blood pressure:
      - Value > 140 and change from baseline > 20

        - Value < 90 and change from baseline < -20
    - Diastolic blood pressure:
        - Value > 90 and change from baseline > 10
        - Value < 55 and change from baseline < -10
- Electrocardiograms (ECGs)
    - Absolute and change from baseline values
    - Marked abnormalities defined by the below categories:
        - QT interval corrected using Fridericia's formula (QTcF):
            - 450 -< 480 msec
            - 480 -< 500 msec
            - ≥ 500 msec
            - 30 < change from baseline ≤ 60 msec
            - Change from baseline > 60 msec
            - Males: < 450 msec, ≥ 450 msec
            - Females: < 470 msec, ≥ 470 msec
        - PR interval ≥ 200 msec
        - QRS interval ≥ 200 msec

In cases where the QT interval is corrected using Bazett's (QTcB) formula is captured instead of QTcF, then QTcB will be converted to QTcF for analyses.

- Eight-Item Patient Health Questionnaire (PHQ-8) total score
    - Absolute and change from baseline values
    - Shifts from baseline scores
- Electronic Columbia-Suicide Severity Rating Scale (eC-SSRS)
    - Suicidal ideation and suicidal behavior responses by visit
    - Shifts from baseline to postbaseline value for suicidal ideation and suicidal behavior
    - Worst postbaseline value for suicidal ideation and behavior

[redacted]

[redacted]

5.0 Populations for Analyses

The following analysis sets will be used in the summary and analysis of study data:

- **Enrolled population**: All subjects who sign informed consent.
- **Full Analysis Set (FAS)**: All subjects who are randomized. Following the intent-to-treat (ITT) principle, subjects will be analyzed according to the treatment group assigned at randomization. The FAS will be the primary efficacy analysis population.
- **Per Protocol Set (PPS)**: A subset of the FAS who are compliant with study treatment and who do not have any relevant protocol deviations that may impact the coprimary efficacy endpoint assessments. The PPS will be analyzed according to the treatment assigned at randomization. The PPS will be a supportive efficacy analysis population and only the co-primary endpoints will be analyzed using this set.
- **As-treated population**: All randomized subjects who take at least one dose of study treatment. Subjects will be analyzed according to treatment received.

[redacted]

5.1 Relevant Protocol Deviations

Relevant protocol deviations are deviations that can have an impact on the primary efficacy endpoints. The impact of relevant protocol deviations on the primary efficacy results will be assessed by excluding subjects from the FAS to define the PPS in supportive analyses of the primary efficacy endpoints. Relevant protocol deviations to be considered regarding exclusion of subjects from the FAS will include the following:

- Randomized but did not take any study treatment
- No postbaseline PASI or sPGA
- Baseline BSA involvement < 10%
- Baseline PASI score < 12
- Baseline sPGA < 3
- Did not have non-plaque psoriasis at baseline
- Poor compliance to study medication within the first 16 weeks of treatment, <75% compliant with study treatment
- Failure to adhere to prohibited concomitant medication restrictions as described below:

- Investigational drug or placebo taken outside of the study any time between Day 1 and the Week 16 assessment
- Phototherapy within 4 weeks prior to the Week 16 assessment
- Biologic medications (eg, adalimumab, etanercept, infliximab, ustekinumab) any time between Day 1 and the Week 16 assessment
- Oral psoriasis medications any time between Day 1 and the Week 16 assessment
- Oral corticosteroids (unless for the treatment of an adverse event) within 4 weeks prior to the Week 16 assessment
- Topical medications/treatments that could affect psoriasis evaluations within 2 weeks prior to the Week 16 assessment
- Medicated shampoos within 2 weeks prior to the Week 16 assessment

- Subject received treatment that was different than intended treatment at any visit prior to Week 16.

All subjects with relevant protocol deviations will be identified prior to database lock and unblinding of treatment assignment with the exception of subjects where actual treatment received is different than randomized treatment which will be determined after treatment unblinding. Relevant protocol deviations will be summarized by treatment group and deviation category for the FAS population.

6.0 Statistical Analyses

Descriptive summaries and analyses will be presented for data captured throughout the study using the following treatment groups.

During the first 16 weeks of treatment, summaries will be provided for the following treatment groups:

- BMS-986165 6 mg QD
- Apremilast
- Placebo

During the first 24 weeks of treatment, summaries will be provided for the following treatment groups:

- BMS-986165 6 mg QD (subjects continuously taking BMS-986165 6 mg QD)
- Apremilast

Summaries will be provided for subjects who have continuous treatment throughout the entire study with either BMS-986165 or apremilast using the following treatment groups:

- BMS-986165 6 mg QD
- Apremilast

Summaries of all subjects exposed to BMS-986165 6 mg QD will also be provided as applicable. Adverse events will be summarized using exposure-adjusted incidence rates (EAIR) for these summaries.

Please note that there will be no Chinese subjects in the apremilast group due to lack of market approval of the agent in China.

6.1 Efficacy Analyses

All efficacy analyses will be performed using the FAS, unless otherwise specified.

Tests of significance of BMS-986165 6 mg QD vs. placebo for the coprimary endpoints will be two-sided with a significance level of 0.05. Both coprimary endpoints need to demonstrate statistical significance to result in a successful study.

The key secondary endpoints will be tested with a two-sided significance level of 0.025 to compare BMS-986165 6mg QD vs. placebo and vs. apremilast. A hierarchical testing approach will be used for testing of

key secondary endpoints (see Section 6.1.3). Two-sided 95% confidence intervals (CIs) will be provided for all efficacy estimates. [redacted]

Coprimary and key secondary efficacy endpoints will be displayed graphically by treatment group and visit, as applicable. All efficacy endpoints will be summarized descriptively.

Due to small sample sizes in China and Japan and differences in stratification factors, analyses will be stratified by a combination of the stratification factors used in randomization. The stratified levels are provided below:

- US/PRIOR USE/< 90 KG
- US/PRIOR USE/>=90 KG
- US/NAÏVE/<90 KG
- US/NAÏVE/>=90 KG
- ROW/PRIOR USE/< 90 KG
- ROW/PRIOR USE/>=90 KG
- ROW/NAÏVE/<90 KG
- ROW/NAÏVE/>=90 KG
- JAPAN/PRIOR USE
- JAPAN/NAÏVE
- CHINA/PRIOR USE
- CHINA/NAÏVE

If there is not at least one subject per treatment group in each of the China and Japan strata, then these regions will be collapsed together to form the following stratified levels for analyses:

- US/PRIOR USE/< 90 KG
- US/PRIOR USE/>=90 KG
- US/NAÏVE/<90 KG
- US/NAÏVE/>=90 KG
- ROW/PRIOR USE/< 90 KG
- ROW/PRIOR USE/>=90 KG
- ROW/NAÏVE/<90 KG
- ROW/NAÏVE/>=90 KG
- ASIA/PRIOR USE
- ASIA/NAÏVE

6.1.1 Primary Endpoints

6.1.1.1 Primary Analysis

Analysis Model

A stratified Cochran-Mantel-Haenszel (CMH) test will be used to compare the sPGA 0/1 response rates at Week 16 between BMS-986165 6 mg QD and placebo using the stratification factors from IRT specified in Section 6.1. The odds ratio (ratio of odds in BMS-986165 6 mg QD group to the odds in the placebo group) and the corresponding 2-sided 95% CIs and p-values will be reported. Additionally, treatment differences in proportions and the 95% CI for the treatment difference will be provided. The common treatment difference in proportion with the confidence interval, common odds ratio with the confidence interval and p-value will be estimated by the Mantel-Haenszel method consistently.

A similar analysis will be performed to compare the PASI 75 response rates at Week 16 between BMS-986165 6mg QD and placebo.

Imputation Methodology

Non-responder imputation (NRI) will be used for coprimary efficacy endpoints for subjects who:

- Discontinue treatment or study prior to Week 16
- Have missing Week 16 endpoint data for any reason

NRI will be used as the primary method of imputation for the coprimary efficacy endpoints.

6.1.1.2 Sensitivity Analyses

As a method to assess the sensitivity of the primary imputation method for the coprimary endpoints, further imputation methods will be used to impute Week 16 data in subjects who:

- Discontinue treatment or study prior to Week 16
- Have missing Week 16 endpoint data for any reason

The coprimary endpoints will be analyzed using the primary analysis method for each sensitivity imputation method described below:

Last Observation Carried Forward (LOCF)

The last observed post-baseline value will be carried forward and used as the Week 16 value. Subjects without a post-baseline will be considered a nonresponder.

LOCF and NRI

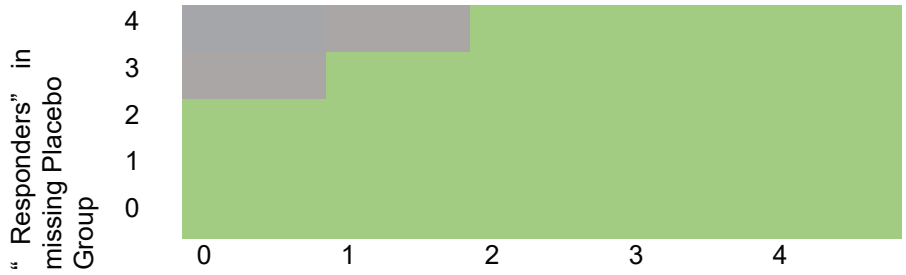
Subjects randomized to the placebo group will have the endpoint value imputed using LOCF (post-baseline). If a placebo subject does not have a post-baseline values, they will be considered a nonresponder. Subjects randomized to BMS-986165 6 mg QD will have their endpoint value imputed using the NRI methodology.

Tipping Point Analysis

Tipping point analysis will further be used to assess the robustness of the primary study results. Different number of events between treatment groups will be assessed until the study conclusion is changed. Each imputed value is initially imputed as a responder. The imputed values in the placebo group will remain as responders while the BMS-986165 imputed values are replaced as a nonresponders one at a time, therefore changing the number of events between groups. Once all imputed values in the BMS-986165 group have been replaced with nonresponders values, the data will reset to where the BMS-986165 group imputed values are all responders and one by one the imputed values in the placebo group are replaced with a nonresponder until all placebo are nonresponders and all BMS-986165 are responders. Furthermore, every pair of imputations between the placebo and BMS-986165 groups will be assessed similarly, creating a matrix of possible patterns.

At each iteration, the statistical analysis is assessed, and the direction of the analysis is recorded. A graph will be provided to identify at what point to which the difference in events cause a direction shift in study results. The tipping point analyses will be based on a two-sample chi-square test approach rather than a stratified CMH test approach for simplicity of the analysis.

Figure 1: Example of Tipping Point Analysis Direction Boundary



"Responders" in missing BMS-986165 Group

Figure 1 represents an example of the tipping-point analysis for all subjects with missing primary efficacy endpoint (responder/non-responder) in BMS-986165 group (N=5) and placebo group (n=5). Gray cells represent pairs where the statistical analysis resulted in non-significance. Green cells represent pairs where the statistical analysis resulted in significant difference between groups. The tipping-point boundary is where the green cells become gray.

Multiple Imputation

Multiple imputation (MI) will be used for sensitivity analyses for each of the coprimary efficacy endpoints, PASI 75 response at Week 16 and sPGA 0/1 response at Week 16. Multiple imputation of missing data for PASI 75 response and sPGA 0/1 response will be performed by fully conditional specification (FCS) using PROC MI in SAS with the FCS option. The FCS MI specifies the multivariate imputation model on a variable-by-variable basis by a set of conditional densities. For each missing value at Week *X*, the FCS MI will generate values from a conditional distribution for the missing data given the other data prior to Week *X*. The FCS MI model will include treatment group and stratification factors for randomization. A total of 1000 imputed (complete) datasets will be generated for the MI analysis. The same test procedure used for the main analysis (CMH test), will be used to analyze the responder endpoint from each imputed dataset. SAS PROC MIANALYZE will be used to pool the results from the CMH tests and generate an overall result by Rubin's rules (1987).

6.1.1.3 Supportive Analyses

Per Protocol Population Analysis

The coprimary endpoints will be analyzed using the PPS using the primary analysis methodology and primary imputation method.

6.1.2 Key Secondary Endpoints

6.1.2.1 Binary Endpoints

Analysis Model

CMH tests will be used to compare response rates between either BMS-986165 6mg QD and placebo or apremilast using the stratification factors from IRT specified in Section 6.1. The odds ratio (ratio of odds in BMS-986165 6 mg QD group to the odds in the placebo or apremilast group) and the corresponding 2-sided 95% CIs and p-values will be reported. Additionally, treatment differences in proportions and the 95% CI for the treatment difference will be provided.

Imputation Methodology for Week 16 Endpoints

NRI will be used for key secondary efficacy endpoints for subjects who:

- Discontinue treatment or study prior to Week 16
- Have missing Week 16 endpoint data for any reason

Imputation Methodology for Week 24 Endpoints (for comparison of BMS-986165 6 mg QD to apremilast)

NRI will be used for key secondary efficacy endpoints for subjects who:

- Discontinue treatment or study prior to Week 24
- Have missing Week 24 endpoint data for any reason

Imputation Methodology for Week 52 Endpoints (for comparison of BMS-986165 6 mg QD to apremilast)

The key secondary endpoints included in this analysis are the following:

- sPGA 0/1 response with at least a 2-point improvement from baseline at Week 52 and at Week 24

- PASI 75 response at Week 52 and at Week 24
- PASI 90 response at Week 52 and at Week 24

Subjects must meet response criteria at both Weeks 24 and 52 in order to be considered a responder. Those who do not meet response criteria at both weeks will be considered a nonresponder. NRI will be used for the above key secondary endpoints for subjects who:

- Discontinue treatment or study prior to Week 52
- Have missing Week 52 or Week 24 endpoint data
- Switch from apremilast to BMS-986165 at Week 24 for Week 24 PASI 50 nonresponders

6.1.2.2 Continuous Endpoints

Analysis Model

Analysis of covariance (ANCOVA) models will be used to compare the change between BMS-986165 6 mg QD and placebo or between BMS-986165 6 mg QD and apremilast at Week 16. Treatment group will be included in the model as well as the stratification factors from IRT specified in Section 6.1 and will be considered as fixed effects. The baseline value will be added into the model as a covariate. The adjusted least-squares (LS) means as well as the treatment differences based on LS means and the corresponding 2-sided 95% CIs will be provided for the difference between BMS-986165 6 mg QD and placebo or apremilast.

Imputation Methodology for Week 16 Endpoints

For continuous key secondary endpoints, a modified baseline observation carried forward (mBOCF) approach will be used for missing data. The baseline observation will be carried forward for subjects who discontinue study treatment prior to Week 16 due to:

- Lack of efficacy
- AEs

Subjects who discontinue study treatment prior to Week 16 for other reasons or who have a missing Week 16 value will have the last valid observation carried forward (including the baseline value as applicable).

Subjects with a missing baseline value will be excluded from the analysis for the change from baseline endpoint.

[redacted]

[redacted]

[redacted]

6.1.4 Subgroup Analyses

Subgroup analyses will be performed on the coprimary endpoints using the FAS. The primary imputation method will be applied for these analyses. The CMH test using the stratification factors from IRT will be the analysis method used. The following subgroups will be considered:

- Geographic region (U.S., Japan, China, Rest of World)
- Country
- Sex (male, female)
- Age group (<65 y, ≥65 y)
- Body weight (<90 kg, ≥90 kg) – from case report form
- Race
- Prior biologic use for psoriasis, psoriatic arthritis and other inflammatory diseases only (yes, no) – from case report form
- Prior systemic treatment use (yes/no)
- Prior phototherapy use (yes/no)
- sPGA (3, 4)
- PASI score (≤20, >20)
- BSA involvement (10-20, >20)
- Duration of disease (< 10 y, ≥ 10 y)
- Age at disease onset (<18, 18-39, ≥40)

[redacted]

[redacted]

[redacted]

6.2 Safety

Summaries of safety data will be presented by period and treatment group, as applicable, for the As-treated population.

6.2.1 Adverse Events

Adverse events will be presented for the number and percentage of subjects and the number of events. Treatment-emergent will be provided in listings. Summary tables will be reported in decreasing frequency based on the BMS-986165 column. Counting for frequency analysis will be by subject and not by AEs, and subjects are only counted once for recurring AEs within each system organ class (SOC) or preferred term (PT); however, the number of total AEs per subject, including multiple occurrences of individual AEs, will also be presented. For summaries of severity, AEs will be reported only at their maximum severity in each subject. Subjects with multiple AEs within a particular SOC or PT will be counted under the category of their most severe event within that SOC or PT.

AE (including deaths) dates will be imputed according to algorithms detailed in Section 8.0. The imputed date of AE onset will be used to assess whether AEs should be considered as treatment-emergent and included in the safety summaries. The original, partial dates will be included in data listings. No imputation will be performed on missing AE seriousness, severity, or relationship; they will be reported as missing.

AEs will be included in a period if the start date of the AE is after the first dispensation date within a period.

An overall summary for the following categories will be presented:

- Deaths
- SAEs
- Related SAEs

- AEs
- Related AEs
- Discontinued treatment due to AEs

The following summaries will also be provided for the following:

- TEAEs by SOC and PT
- Treatment-related TEAEs by SOC and PT
- TEAEs by PT reported in ≥ 5% of subjects
- TEAEs by PT reported in ≥ 1% of subjects
- Treatment-related TEAEs by PT reported in ≥ 5% of subjects
- TEAEs categorized by severity by SOC and PT
- Exposure-adjusted incidence rate (EAIR) for TEAEs by SOC and PT – EAIR is defined in Section 8.1 of the SAP

6.2.1.1 Adverse Events of Interest (AEI) and Other Important AEs

Summaries for treatment-emergent AEIs events will be provided by PT for each AEI category:

- Skin-related events
- Infection events
- Malignancy events

Creatine kinase (CK) elevation for CK elevation > 2.5 upper limit of normal will be summarized as CTCAE grade 2 or higher in the clinical laboratory summaries.

Additional information collected for some events as part of the clinical safety program and adjudicated events will also be summarized.

6.2.1.2 Serious Adverse Events

Summaries for treatment-emergent SAEs will be provided for the following:

- Treatment-emergent SAEs by SOC and PT

6.2.1.3 Adverse Events Leading to Discontinuation of Study Treatment or Study Treatment Interruption

Summaries for TEAEs leading to discontinuation of study treatment will be provided for the following:

- TEAEs by SOC and PT

Summaries for TEAEs leading to study treatment interruption will be provided for the following:

- TEAEs by SOC and PT

6.2.2 Deaths

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

6.2.3 Clinical Laboratory Data

Laboratory parameters will be summarized using the International System (SI) of Units and US conventional units. Data will be summarized by time point, as applicable. The following summaries will be provided for each parameter:

- Absolute and change from baseline values for continuous parameters

- Number and percentage of subjects for the following:
    - Maximum postbaseline CTCAE grade for each applicable laboratory parameter through Week 16
    - Shifts from baseline based on maximum postbaseline CTCAE grade through Week 16
- Drug-induced Liver Injury (DILI) and Hy's Law summaries

6.2.4 Vital Signs and Physical Findings

Vital signs, including weight, will be summarized by time point, as applicable. The following summaries will be provided for each parameter:

- Absolute and change from baseline values
- Number and percentage of subjects for the following:
    - Marked abnormality for each abnormality category as defined in Section 4.2

6.2.5 ECGs

ECG parameters will be summarized by time point, as applicable. The following summaries will be provided for each parameter:

- Absolute and change from baseline values
- Number and percentage of subjects for the following:
    - Marked abnormality for each abnormality category as defined in Section 4.2

6.2.6 Other Safety Data

6.2.6.1 PHQ-8

PHQ-8 total score will be summarized by time point, as applicable. The following summaries will be provided:

- Absolute and change from baseline values
- Number and percentage of subjects:
    - Shifts from none, mild, moderate, moderately severe and severe scores at baseline and at each time point

6.2.6.2 eC-SSRS

Suicidal ideation and behavior individual item responses will be summarized by time point, as applicable. The following summaries will be provided:

- Number and percentage of subjects with positive responses on suicidal ideation and/or suicidal behavior questions for each question and overall all questions within suicidal ideation and suicidal behavior
- Shifts from baseline based on maximum postbaseline response through Week 16
- Worst postbaseline value for suicidal ideation and behavior through Week 16

6.3 General Methodology

The following standards/ methods will be used:

- Statistical package(s) planned to be used
    - All analyses will use SAS version 9.4 or higher.
- Standard summary statistics for continuous and categorical variables:

- Unless otherwise noted, categorical variables will be summarized using counts and percentages. Percentages will be rounded to one decimal place, except 100% will be displayed without any decimal places and percentages will not be displayed for zero counts.
- Continuous variables will be summarized using the number of observations (n), mean, standard deviation (SD), median, minimum and maximum. The median, minimum and maximum values will be displayed to the same level of precision as the raw data, the mean to a further decimal place and the SD to two additional decimal places.
- Variables will be summarized by period, treatment group, and time point, as applicable.

6.3.1 Subject Populations and Disposition

The number of subjects enrolled/screened and the number and percentage of subjects randomized, treated, and in each analysis population will be presented. The number and percentage of subjects randomized in each region, country, and site will be presented.

Additionally, the following summaries will be provided for the FAS by treatment group and overall:

- Number and percentage of subjects who completed 16 weeks of treatment
- Number and percentage of subjects who discontinued treatment prior to Week 16 and reason for treatment discontinuation
- Number and percentage of subject who completed 24 weeks of treatment, who discontinued treatment prior to Week 24 and post Week 16, and those who discontinued at any time prior to week 24 and reason for treatment discontinuation. Denominator will be the number of subjects in the FAS.
- Number and percentages of subjects who completed 52 weeks of treatment, who discontinued treatment prior to Week 52 and post Week 24, and who discontinued treatment at any time and reason for treatment discontinuation. Denominator will be the number of subjects in the FAS.

6.3.2 Demographic and Baseline Characteristics

Demographic and baseline characteristics will be summarized by treatment group for the FAS. Demographic characteristics include the following:

- Sex
- Race
- Ethnicity
- Age (in years, at time of signing informed consent) and age category (<65 vs ≥65)
- Weight (in kg, at baseline) and weight category (≥90 kg, <90 kg)
- Body mass index (BMI in kg/m^2, at baseline)
- Geographic region (U.S., Japan, China, Rest of World)
- Prior biologic use for psoriasis, psoriatic arthritis and other inflammatory diseases only (yes, no)
- Reason for discontinuation of prior biologic use
- Prior systemic treatment use (yes/no)
- Prior phototherapy use (yes/no)
- Stratification factors obtained from IRT: geographic region by prior biologic use and by body weight
- Stratification factors obtained from database: geographic region by prior biologic use and by body weight
- sPGA (3, 4)
- PASI score (≤20, >20)
- BSA involvement (10-20, >20)
- Duration of disease (< 10 y, ≥ 10 y)
- Age at disease onset (<18, 18-39, ≥40)

Additional demographics or baseline data may be added to summary tables.

General medical history and medical history related to psoriasis will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). General medical history data will be summarized for each SOC and PT by treatment group and overall for the FAS. Separate tables will be provided for psoriasis medical history.

6.3.3 Prior and Concomitant Medications

Prior and concomitant medications will be coded according to World Health Organization-Drug Dictionary WHO-DD (version at the time of DBL) and will be summarized by Anatomic Therapeutic Classification (ATC) and preferred term (PT) by treatment group for the As-treated population. The number and percentage of subjects using at least one medication and each medication will be displayed by treatment group. Subjects taking more than one medication within the same ATC or PT will be counted once.

Summaries will be provided for prior medications and medications that started prior to first treatment and were ongoing after first treatment start date as well concomitant medications.

Medication dates will be imputed according to algorithms detailed in Section 8.0. The imputed dates will be used to assess whether medications should be included in the summaries as prior or concomitant, however the original, partial dates will be included in data listings.

6.3.3.1 Psoriasis, Psoriatic Arthritis and Other Inflammatory Disease Related Systemic Medications

Prior medications and medications that started prior to first treatment and were ongoing after first treatment start date for systemic biologic and non-biologic medications will be summarized as described above for the number and percentage of subjects using each reported medication.

6.3.3.2 Concomitant High Potency Corticosteroid Use

The number and percentage of subjects using at least one high potency topical corticosteroid at Week 24 will be summarized by treatment group. Additionally, corticosteroids will be summarized by ATC and PT.

6.3.4 Exposure

6.3.4.1 Duration of Treatment

Duration by Group

Overall duration of each treatment, in days, will be calculated for each subject in each treatment group. Subjects are dispensed study treatment at each visit, starting with Day 1 (Week 0) and at each subsequent visit. The date of first dose of study treatment is the Week 0 [redacted] dosing date and is recorded on the eCRF. If this date is missing, then the earliest drug dispensation date will be used. The last dose date is recorded on the end of treatment page in the eCRF. If this date is missing, then the latest date from the [redacted] page or drug accountability return date will be used.

Duration of treatment will be summarized descriptively by randomized treatment group.

BMS-986165:

Subjects randomized to BMS-986165 will have their duration of treatment derived as:

- Date of last dose – date of first dose +1

Placebo:

For subjects randomized to placebo, duration is defined as:

- Placebo = Date of last dose of placebo – date of first dose +1
- BMS-986165 = Date of last dose of BMS-986165 – date of first dose of BMS-986165 +1

Apremilast:

For subjects randomized to apremilast, the Week 24 date will be used as the date of last dose of apremilast and the date of first dose of either apremilast or BMS-986165 for the subsequent period. Subjects who receive apremilast at Week 24 through to Week 52 will be counted as a sum of the first 24 weeks of treatment and the next treatment period.

- Apremilast (Wk 0 through Wk 24) = Week 24 date – date of first dose +1
- Apremilast (Wk 24 through Wk 52) = Date of last dose – Week 24 date
- Total Apremilast = Duration of apremilast Wk 0-24 + duration of apremilast Wk 24-52
- BMS-986165 = Date of last dose of BMS-986165 – Week 24 date

Total apremilast duration of treatment will be displayed as well as BMS-986165 treatment duration. If subject discontinues their study treatment during the initial 24 weeks of apremilast treatment, the date of last dose will be used as the duration of apremilast.

Duration by Period

Overall duration (in days) of each treatment received within each study period will be calculated. Duration within each period is defined for each treatment group as:

$$Last\ dose\ date\ in\ period\ - first\ dose\ date\ in\ period + 1$$

6.3.4.2 Summary of Dosing

The number of doses taken for each subject for each period will be determined using the eCRF drug accountability data and is defined as:

$$Doses\ Taken: (number\ of\ tablets\ dispensed - number\ of\ tablets\ returned)$$

The number of doses taken will be summarized descriptively by treatment group within each period and overall.

6.3.4.3 Compliance

Treatment compliance will be determined from data captured on the Drug Accountability eCRF.

$$Number\ of\ expected\ doses: (date\ of\ next\ visit - date\ of\ current\ visit)x3$$

Treatment compliance will be derived for each period. Compliance is defined as:

$$\left(\frac{Number\ of\ doses\ taken}{Number\ of\ expected\ doses}\right)x100$$

Period compliance will be calculated by summing over all visits within the period using descriptive statistics by treatment group. The number and percentage of subjects with <75%, 75% to 100%, and >100% compliance will be provided by treatment group for each period. If a subject does not return the container, then this dispensation event will be excluded in the calculation for the total number of doses taken, total number of expected doses, and total compliance.

[redacted]

[redacted]

6.6 Statistical Impacts Due to COVID-19

6.6.1 Impact on Efficacy Endpoints

There are no COVID-19-related impacts to the Week 16 and Week 24 endpoints as all subjects remaining in the trial completed these visits prior to COVID-19 restrictions. Key secondary efficacy endpoints involving later visits may be impacted by COVID-19 in the form of missing data and/or a treatment assignment defaulting to BMS-986165. If a subject has missing data or is switched to BMS-986165 from a different treatment due to COVID-19 during the Week 24 through Week 52 period, the subject will be excluded from the analysis.

Key secondary efficacy endpoints that may be impacted include the following:

- sPGA 0/1 response at Week 52 and at Week 24
- PASI 75 response at Week 52 and at Week 24
- PASI 90 response at Week 52 and at Week 24

A sensitivity analysis will also be performed for the above endpoints using the last observed value. The data handling rules to be used are as follows for subjects with missing data or for those subjects who switch to BMS-986165 from a different treatment due to COVID-19:

- If the subject is missing the Week 52 response value (ie, sPGA 0/1, PASI 75, PASI 90), then the last observed response value during the Week 24 through Week 52 period will be carried forward to the Week 52 value
- If the subject is switched to BMS-986165 from a different treatment due to COVID-19, then the last observed response value prior to the switch will be carried forward to the Week 52 value

Censoring rules will be applied to the time to first loss endpoints evaluated for the maintenance of efficacy of BMS-986165 through Week 52 in subjects who were originally randomized to BMS-986165 and impacted by COVID-19-related issues post Week 24:

- Subject has missing visits (ie, efficacy assessments not performed) after Week 24 and assuming loss of response was not observed prior to missing visits:

- If the subject returns at a later visit and is found to have had a loss of response, midpoint imputation will be used to determine time to first loss (ie, midpoint of the censoring interval).
- If the subject returns at a later visit and has not had a loss of response, then the subject will continue to be evaluated for time to first loss.
- If the subject is discontinued from the study treatment, then the subject will be censored at the time of study treatment discontinuation.

Time to first loss endpoints include:

- Time to first loss of PASI 75 among subjects that are PASI 75 responders at Week 24
- Time to first loss of sPGA 0/1 among subjects that are sPGA 0/1 responders at Week 24 where loss of sPGA 0/1 is defined as an sPGA score ≥ 2 (sPGA scores are rounded to the nearest whole number)

6.6.2 Impact on Safety Endpoints

No modifications will be made for the analysis of safety data due to COVID-19-related issues. There will not be any additional safety data handling considerations for missing data or treatment interruptions due to COVID-19.

7.0 Sequence of Planned Analyses

7.1 Interim Analyses

No interim analysis is planned for this study.

7.2 Final Analyses and Reporting

All final, planned analyses identified in this statistical analysis plan will be performed only after the last subject has completed the study and the database has been locked. The randomization codes for all subjects will not be unblinded until after the database has been locked.

Analyses of the palmoplantar pp-PGA 0/1 endpoint are provided for descriptive purposes within the individual CSR due to small sample sizes. A pooled analysis of data from IM011046 and IM011047 will be conducted for the submission.

8.0 Conventions

8.1 General Definitions

The following data definitions and handling conventions will be used for general analysis:

Term	Definition
Study Day	Study day is calculated as: assessment date – date of first dose + 1
Baseline	Unless otherwise stated, Baseline is defined as the measurement at the randomization visit (Week 0). If the measurement at the randomization visit is missing, then a prior measurement during the screening period may be used as baseline. Baseline assessments must be performed per protocol and standard of care assessments may not be used for baseline. Baseline for each PSSD question will be calculated as the average value over the 7 days prior to the randomization visit. In the event of missing values in the 7 days prior to the randomization visit, daily

	scores from at least 4 days out of the 7 days can be used to calculate the average score. If > 3 daily scores in the last 7 days are missing for a particular question, then the baseline value for that question will be missing. If any questions are missing baseline after using the described rule, then the baseline value for the subscore using that question and the baseline value for the total score will be missing.
Change from Baseline	Change from baseline is defined as (value at post-baseline visit – value at baseline).
Change in the maximum post-baseline value or change in the worst post-baseline value	Change from baseline in the maximum post-baseline value is defined as highest observed value or grade post-baseline. The change is calculated using this value as the post-baseline value. Change from baseline in the worst post-baseline value is defined as worst observed value post-baseline. The change is calculated using this value as the post-baseline value.
Concomitant and Prior Medication	Prior medications are defined as medications with a stop date prior to the first dose of study treatment. Concomitant medications are defined as any medications ongoing at the start of study treatment or with a start date on or after the first dose date.
End of Study (EOS) Date	The EOS date is the date recorded on the eCRF that a randomized subject either discontinued or completed the study. If the subject is lost to follow-up, the EOS date will be the date of the last visit assessment obtained.

Exposure-adjusted incidence rate (EAIR)	EAIR = 100* 365.25*(total number of subjects with the AE)/total exposure time for the selected AE under each treatment that a subject is exposed. Where total exposure time for each AE within a treatment is calculated as follows:
	
	• If a subject has at least 1 event while on a particular treatment, then the exposure time for that subject and AE on that treatment is:
	  ○ First AE onset date – treatment start date (of that particular treatment) +1
	• If a subject does not have an event, exposure time for that AE is:
	  ○ Treatment stop date (of that particular treatment) – Treatment start date (of that particular treatment) + 1 + 30 days (if subject discontinued or subject completed Period 3 and is not rolling into IM011075 study)
	  ○ Treatment stop date (of that particular treatment) – Treatment start date (of that particular treatment) + 1 (if subject completed Period 3 and is rolling into IM011075 study)
	• Total exposure time = sum of exposure time for each AE within a treatment

First Dose Date – Study	The date a subject received their first dose on Day 1 as recorded in the eCRF Week 0 [redacted] dosing date or the earliest drug dispensation date.

Last Dose Date – Study	The date of last recorded dose on the eCRF for a randomized subject.
First Dose Date – Period	The date a subject received their first dose as recorded in the eCRF Week 0 [redacted] dosing date or the earliest drug dispensation date for Treatment Period 1 and the earliest drug dispensation date for Treatment Periods 2 and 3.
Last Dose Date – Period	The date of the last visit in the periods – 1. If a subject prematurely discontinues study treatment within a period, the date of last recorded dose on the eCRF will be used as the last dose date for the period.
Percent Change from Baseline	Percent change from baseline is defined as ([value at post-baseline visit – value at baseline]/value at baseline) x 100. If the baseline value is 0 and the post-baseline value is also 0, then the percent change from baseline is set to 0. If the baseline value is 0 and the post-baseline is >0, then the percent change from baseline value will be missing.

8.2 Calculation of Key Measures

The following efficacy assessments will be used to assess subjects' disease activity and severity during the study. Outcomes are reported via an eCOA tool at various times throughout the study as described in the protocol Schedule of Activities. At study visits, assessments by the investigator or subjects and results/responses will be reported directly into the eCOA tool at the time of the visit. The tool will open assessments in a sequential manner, meaning that the full assessment is to be completed prior to moving forward to the next assessment. This limits the possibility of partially missing data. Also, as investigators/subjects are prompted to enter data for each assessment for the visit, the possibility of a full assessment being missing is also negated.

Scoring of assessments where validated algorithms are not required will be derived in SAS datasets.

Scoring of assessments where validated scoring tools are required, licenses for these tools will be purchased and used for scoring prior to incorporating into the SAS datasets.

8.2.1 Investigator-Administered Assessments

Assessments will be performed by a qualified physician or dermatologist or trained designee who is experienced in the assessment of psoriasis patients. To limit variability, every effort will be made so that the same individual conducts the assessment at all subsequent visits.

8.2.1.1 static Physician's Global assessment (sPGA)

The sPGA is a 5-point scale of an average assessment of all psoriatic lesions based on erythema, scale, and induration. The sPGA measure determines psoriasis severity at a single point in time (without taking into account the baseline disease condition) as clear (0), almost clear (1), mild (2), moderate (3), or severe (4). A higher score equates to higher severity of disease.

The individual scores at each visit for erythema (E), induration (I) and scaling (S) will be captured via the eCOA system. Scores will range from 0 to 4. A total score will also be computed based on the average of the 3 characteristic scores.

$$Total\ average\ score = \frac{E + I + S}{3}$$

The total average score will be calculated in the eCOA system. The average score will be rounded to the nearest whole number. For example, if the total average score is ≤1.49 the score will be rounded to 1. If

the score is ≥1.5 the score will be rounded to 2. The primary endpoint is derived from the total average score.

sPGA 0 is derived as the binary indicator for sPGA from the calculation above equal to 0 or not;

sPGA 0/1 is derived as the binary indicator for sPGA from the calculation above equal is less than 2 or not;

[redacted]

All individual scores and total average score assessed at each week throughout the study will be transferred to PRA for analysis. The endpoint derivations will be performed in the analysis datasets.

8.2.1.2 Psoriasis Area and Severity Index (PASI)

The PASI is a measure of the average redness, thickness, and scaliness of psoriatic skin lesions (each graded on a 0–4 scale; 0 = none to 4 = very severe), weighted by the area of involvement (head, arms, trunk to groin, and legs to top of buttocks). The PASI produces a numeric score that can range from 0 to 72, with higher PASI scores denoting more severe disease activity. The PASI includes multiple subscores and a final total score that will be provided by the eCOA system. Individual plaque characteristic rating scores are provided for each body region as well as the weighted score. Additionally, the degree of involvement of each body region is assessed and that score is multiplied by the weighted plaque characteristic score for a final score for each body region. The total PASI score is a sum of the 4 body regions: Head, Upper Extremities, Trunk and Lower Extremities.

The PASI Total score will be used to assess response to treatment. The percent change from baseline will be calculated at each visit. The PASI 75 endpoint is the proportion of subjects who experience at least a 75% improvement in PASI score as compared with the baseline value.

1 = $If\ \left(\frac{Baseline\ PASI - Visit\ PASI}{Baseline\ PASI}\right) x100 \geq 75\ then\ subject\ is\ a\ PASI\ 75\ responder$

0 = otherwise

The PASI 50, PASI 90, and PASI 100 are defined similarly. The endpoint derivations will be performed in the analysis datasets.

8.2.1.3 Body Surface Area (BSA)

Measurement of psoriasis BSA involvement is estimated using the handprint method with the size of a subject's handprint (including fingers and thumb) representing 1% of BSA involved. The total BSA = 100% with breakdown by body region as follows:

- Head and neck = 10% (10 handprints),
- Upper extremities = 20% (20 handprints),
- Trunk including axillae and groin = 30% (30 handprints),
- Lower extremities including buttocks = 40% (40 handprints).

The Total BSA is the sum of each body region and is assessed at each visit and recorded in the eCOA system.

The product of BSA and sPGA will be calculated. At baseline, baseline BSA will be multiplied by baseline sPGA score. The derivation will be performed at each subsequent visit.

8.2.1.4 scalp specific Physician's Global Assessment (ss-PGA)

The scalp specific assessment will only be performed in subjects with scalp involvement. If there is evidence of scalp involvement, scalp lesions are evaluated in terms of clinical signs of redness, thickness, and scaliness and scored on the following 5-point ss-PGA scale:

0 = absence of disease, 1 = very mild disease, 2 = mild disease, 3 = moderate disease, 4 = severe disease.

The ss-PGA is assessed at each visit throughout the study in subjects that have evidence of scalp psoriasis at baseline. The score will be collected in the eCOA system.

[redacted]

8.2.1.6 Physician's Global Assessment-Fingernails (PGA-F)

In this assessment, fingernail psoriasis is evaluated. The PGA-F will be performed at baseline. If a subject shows evidence of psoriatic fingernail involvement, the assessment will be performed at each subsequent visit to assess severity and improvement over time. Only subjects with PGA-F at baseline will be assessed throughout the study. The overall condition of the fingernails is rated on a 5-point scale:

0 = clear, 1 = minimal, 2 = mild, 3 = moderate, and 4 = severe

The rating score will be collected in the eCOA system.

[redacted]

8.2.1.8 Palmoplantar PGA (pp-PGA)

This measure will be used for subjects with palmoplantar (finger and toe surfaces) involvement at baseline. Only subjects with baseline palmoplantar involvement will continue to have these assessments at each subsequent visit throughout the study. The pp-PGA uses a 5-point (0-4) overall severity scale:

0 = clear; 1 = almost clear; 2 = mild; 3 = moderate; and 4 = severe.

Scores are collected at each visit and entered in the eCOA system.

[redacted]

8.2.2 Subject-Reported Assessments

8.2.2.1 Psoriasis Symptoms and Signs Diary (PSSD)

The PSSD is an 11-item subject-reported instrument that assesses severity of symptoms and subject-observed signs commonly associated in plaque psoriasis. It has been shown to be reliable and valid in measuring symptoms and signs of subjects with moderate-to-severe plaque psoriasis in the clinic and has strong psychometric properties in assessing treatment effects in clinical trials. The PSSD assesses severity of 5 symptoms (itch, pain, stinging, burning, skin tightness) and 6 subject-observed signs (skin dryness, cracking, scaling, shedding or flaking, redness, bleeding) using 0–10 numerical ratings. The severity of each item is rated on an 11-point numeric rating scale ranging from 0 (absent) to 10 (worst imaginable).

The following questions in the instrument are included in the symptom score: Q1, Q4, Q9, Q10, Q11

The following questions in the instrument are included in the sign score: Q2, Q3, Q5, Q6, Q7, Q8

Two versions of the PSSD were developed, one with a 24-hour recall period (PSSD-24h) and one with a 7-day recall period. The PSSD-24h will be administered daily in this trial to avoid recall bias with a longer recall period. Individual scores to each question are collected daily in the PSSD. For visit PSSD scoring, the daily scores (with 24-h recall periods) over the prior 7 day will be used and the average score to each of the 11 questions will be used as the score at that visit. In case missing data arise during the 7 days prior to the visit, daily scores of at least 4 days out of the 7 can be used. If >3 scores are missing, the average score will be missing. Baseline PSSD score is calculated based on the daily diary collected data during the screening period. Baseline for each PSSD question will be calculated as the average value over the 7 days

prior to the randomization visit. In the event of missing values in the 7 days prior to the randomization visit, daily scores from at least 4 days out of the 7 days can be used to calculate the average score. If > 3 daily scores in the last 7 days are missing for a particular question, then the baseline value for that question will be missing. If any questions are missing baseline after using the described rule, then the baseline value for the subscore using that question and the baseline value for the total score will be missing.

A symptom score will be derived by averaging the 5 questions included in the symptom score and multiplying by 10. To obtain a symptom score on a given day, responses to at least 2 of the 5 questions must be available. If 3 or more questions are missing, the symptom score is considered missing.

A sign score will be derived by averaging the 6 questions included in the sign score and multiplying by 10. Responses to at least 3 of the 6 questions must be available in order to obtain a sign score for a given day. If more than 3 questions are missing, the sign score is considered missing.

Both scores range from 0-100, where 0 representing the least severe symptom/sign and 100 the most severe. A total PSSD score with range 0-100 will be derived from taking the average of the symptom and sign scores.

[redacted]

8.2.2.4 Dermatology Life Quality Index (DLQI)

The DLQI is a subject-reported quality of life index which consists of 10 questions concerning symptoms and feelings, daily activities, leisure, work, school, personal relationships, and treatment during the last week. Each question is scored on a scale of 0 to 3 where 0="not at all", 1="a little", 2="a lot", or 3="very much". The scores are summed, giving a range from 0 (no impairment of life quality) to 30 (maximum impairment). Question 7 includes 2 questions, if the subject answers 'Yes' to Q7, the score given is a 3. If the subject answers 'No' to Q7, they are asked the second question where a score of 0='not at all', 1='a little', or 2='a lot' is given. Certain questions include an option for not relevant. When scoring, any questions deemed 'not relevant' will take on a value of 0.

. Interpretation of DLQI scores is as follows:

- 0-1 = no effect at all on patient's life
- 2-5 = small effect on patient's life
- 6-10 = moderate effect on patient's life
- 11-20 = very large effect on patient's life
- 21-30 = extremely large effect on patient's life

Individual scores for each question will be provided by the eCOA system. The DLQI score will be derived in the analysis datasets.

[redacted]

[redacted]

8.2.2.11 Psoriatic Arthritis Screening and Evaluation (PASE) Questionnaire

The PASE questionnaire will be administered at Screening in subjects with peripheral joint complaints. The PASE questionnaire is a self-administered tool that is used to screen for psoriatic arthritis among subjects who have psoriasis. The PASE questionnaire consists of 15 questions subdivided into 7 questions focusing on symptoms of psoriatic arthritis, and 8 questions focusing on the impact of psoriatic arthritis on function. Each question is scored on a 1 to 5 scale, with a maximum symptom sub-scale score of 35, a maximum function sub-scale score of 40 giving a total maximum score of 75. A score of 47 or above has been shown to distinguish between psoriatic arthritis and non-psoriatic arthritis. This questionnaire is only intended to be a screening tool for psoriatic arthritis and does not replace a comprehensive musculoskeletal exam performed by a rheumatologist. This information will not be summarized.

Individual score to each of the 15 questions will be collected in the eCOA system. The sub-scale scores and total score will be derived in the analysis datasets.

8.2.2.12 Eight-Item Patient Health Questionnaire (PHQ-8)

The PHQ-8 is established as a valid diagnostic and severity measure for depressive disorders in large clinical studies. Each of the 8 questions is based on a 2-week recall and scored on a scale of 0 to 3 by a tick box as: 0=Not at All, 1=Several Days, 2=More than Half the Days, and 3=Nearly Every Day. A PHQ-8 score is derived by summing the scores for the 8 questions. The total PHQ-8 score ranges from 0-24. Scoring interpretation is as follows:

- 0-4 = no significant depressive symptoms
- 5-9 = mild depressive symptoms
- 10-14 = moderate depressive symptoms
- 15-19 = moderately severe depressive symptoms
- 20-24 = severe depressive symptoms

Response to each individual question is collected in the eCOA system. The total score will be derived in the analysis datasets.

8.2.2.13 electronic Columbia-Suicide Severity Rating Scale (eC-SSRS)

The electronic Columbia-Suicide Severity Rating Scale (eC-SSRS) is a computer-automated, patient-reported version of the C-SSRS instrument that defines 11 categories of suicidal ideation and behavior (SIB) events. Categories and definitions are provided in Appendix 23 of the protocol. The categories are as follows:

- Suicidal ideation
  1. Wish to be dead
  2. Non-specific active suicidal thoughts
  3. Active suicidal ideation with any methods without intent to act
  4. Active suicidal ideation with some intent to act, without specific plan
  5. Active suicidal ideation with specific plan and intent
- Suicidal behavior
  1. Preparatory acts or behavior
  2. Aborted attempt
  3. Interrupted attempt
  4. Actual attempt (Non-fatal)
  5. Completed suicide
- Self-injurious behavior, no suicidal intent

8.3 Missing, Unknown, or Partial Dates

Start Date		Stop Date						
		Complete: *yyyymmdd*		Partial: *yyyymm*		Partial: *yyyy*		Missing/ Ongoing
		<1st dose	≥1st dose	<1st dose *yyyymm*	≥1st dose *yyyymm*	<1st dose *yyyy*	≥1st dose *yyyy*	
Partial: *yyyymm*	= 1st dose *yyyymm*	2	1	n/a	1	n/a	1	1
	≠ 1st dose *yyyymm*		2	2	2	2	2	2
Partial: *yyyy*	= 1st dose *yyyy*	3	1	3	1	n/a	1	1

	≠ 1st dose *yyyy*		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1 = Impute as the date of first dose
2 = Impute as the first of the month
3 = Impute as January 1 of the year
4 = Impute as January 1 of the stop year
Note: If the start date imputation leads to a start date that is after the stop date, then there is a data error and the start date is not imputed.

Imputation rules for partial or missing stop dates:

1. Initial imputation
    a. For partial stop date "mmyyyy", impute the last of the month.
    b. For partial stop date "yyyy", impute December 31 of the year.
    c. For completely missing stop date, do not impute.
2. If the stop date imputation leads to a stop date that is after the death date, then impute the stop date as the death date.
3. If the stop date imputation leads to a stop date that is before the start date, then there is a data error and the stop date is not imputed.

Imputation rules for partial or missing death dates:

1. If death year and month are available but day is missing:
    a. If "mmyyyy" for last contact date = "mmyyyy" for death date, set death date to the day after the last contact date.
    b. If "mmyyyy" for last contact date < "mmyyyy" for death date, set death date to the first day of the death month.
    c. If "mmyyyy" for last contact date > "mmyyyy" for death date, data error and do not impute.

If both month and day are missing for death date or a death date is totally missing, set death date to the day after the last contact date.

8.4 Study Periods

Period 1 = Week 0 to Week 16 visit date

Period 2 = Week 16 visit date +1 to Week 24 visit date

Period 3 = Week 24 visit date +1 to Week 52 visit date

Follow-up = 4 week follow-up period

8.5 Day Ranges for Analysis Visits

Below are the day ranges for the analysis visit definitions. If more than one visit occurs within an analysis visit, then the visit that is closest to the target date should be used for analysis.

Period **Week**	**Target Day**	**Day Range**
Baseline		Screening, 1
Period 1		
Week 1	8	2, 11
Week 2	15	12, 18
Week 4	29	19, 43
Week 8	57	44, 71
Week 12	85	72, 99
Week 16	113	100, 127 (or Week 16 drug dispense date)
Period 2		
Week 20	141	1st day after Week 16 drug dispense date, 155

Period Week	Target Day	Day Range
Week 24	169	156, 183 (or Week 24 drug dispense date)
Period 3		
Week 28	197	1st day after Week 24 drug dispense date, 211
Week 32	225	212, 239
Week 36	253	240, 267
Week 40	281	268, 295
Week 44	309	296, 323
Week 48	337	324, 351
Week 52	365	352, last visit date prior to Safety Follow-up
Safety Follow-up	393	Safety Follow-up visit

9.0 References

- Campbell G, Pennello G, Yue L. Missing data in the regulation of medical devices. Journal of Biopharmaceutical Statistics 2011;21:180-195.
- Yan X, Lee S, Li N. Missing data handling methods in medical device clinical trials. Journal of Biopharmaceutical Statistics 2009;19:1085-1098.
- [redacted]
- Kenward MG, Roger J. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997;53(3):983-997.
- Rubin, DB (1987). Multiple Imputation for Nonresponse in Surveys. New York, J. Wiley & Sons.

10.0 Document History

Version Number	Version Date	Summary of Changes
1.0 - Original Document	09-Jun-2019	Not applicable
2.0 – Amendment 01	30-Sep-2020	Revisions provided below.

Revisions for Amendment 01: In addition to the revisions specified below, there were some minor typographical and formatting changes made. Additions are noted by bold text. Removals are noted by strikethrough.

SAP Section	Revised Text	Rationale for Change
2.1 Study Design	**Week 16 activities** The coprimary endpoints, static Physician Global Assessment (sPGA) 0/1 and Psoriasis Area and Severity Index (PASI) 75 will be assessed at Week 16. Subjects receiving placebo will be switched in a blinded manner to BMS-986165 6 mg QD at Week 16. Subjects who are randomized to BMS-986165 6 mg QD or apremilast will continue on their assigned treatment regimen **in a blinded manner**.	Minor edit to clarify subjects continue on assigned treatment in a blinded fashion.
2.4 Unblinding Information	~~No other unblinding prior to study completion and database lock is planned~~. **Additionally, bioanalytical scientists involved in the processing of bioanalytical samples will be unblinded to randomized treatment assignments to minimize unnecessary sample bioanalysis of subjects who are on placebo.**	Added clarification that bioanalytical staff will be unblinded to facilitate bioanalytical sample processing.
2.5 Changes in Statistical Considerations from the Protocol	~~There are no changes in statistical considerations from the protocol at this time.~~ The following is a list of the important changes in the SAP from the Statistical Considerations section in the protocol: • **The hierarchical testing order for the key secondary endpoints has been updated and two separate hierarchies have been provided, one for US submission and one for ex-US submission (see Tables 1 and 2 in Sec. 6.1.3). The hierarchies presented here supersede the one that is in the protocol.** • **Key secondary endpoints were added for the comparison of BMS-986165 to apremilast for sPGA 0/1, PASI 75, and PASI 90 at Week 24.** • **Imputation methods were updated to remove the prohibited medication/therapy criteria for binary and continuous endpoints.**	The definition for the Full Analysis Set was modified to include all randomized subjects as this is more representative of the ITT population. Noted that the relevant deviation list was updated from the list in the protocol.

	• **The Full Analysis Set was changed from "all subjects who were randomized to receive assigned study treatment" in the protocol to "all subjects who are randomized".** • **The list of relevant protocol deviations provided in the protocol was updated for determination of the Per Protocol population.** • **Logistic regression analyses for binary endpoints were removed as these are similar to the CMH analyses.**	Additional Week 24 endpoints were added. Imputation methods were updated to remove prohibited medication/ therapy criteria. Hierarchical testing order of key secondary endpoints were updated and accounting of regional regulatory needs. Removed logistic regression analyses.
4.1.2 Secondary Endpoints 4.1.2.1 Key Secondary Endpoints for Comparisons to Placebo	The key secondary efficacy endpoints for BMS-986165 compared to placebo at Week 16 are defined as: • PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in PASI score • sPGA 0 response assessed as a proportion of subjects with an sPGA score of 0 • Scalp specific Physician's Global Assessment (ss-PGA) 0/1 assessed as a proportion of subjects with a ss-PGA score of 0 or 1 with at least a 2-point improvement from baseline among subjects with a baseline ss-PGA score ≥3 • PSSD symptom score of 0 assessed as a proportion of subjects with a PSSD symptom score of 0 among subjects with a baseline PSSD symptom score ≥1 • PASI 100 response assessed as a proportion of subjects who achieve a 100% improvement from baseline in PASI score • Physician Global Assessment-Fingernails (PGA-F) 0/1 assessed as a proportion of subjects with a PGA-F score of 0 or 1 with at least a 2-point improvement from baseline among subjects with a baseline PGA-F score ≥3 • ~~Change from baseline in Psoriasis Symptoms and Signs Diary (PSSD) symptom score~~	Updated key secondary endpoints compared to placebo and re-ordered.

	• DLQI 0/1 assessed as a proportion of subjects with a DLQI score of 0 or 1 among subjects with a baseline DLQI score ≥2 (Ex-US submission only) • ~~Palmoplantar Physician's Global Assessment (pp-PGA) 0/1 assessed as a proportion of subjects with a pp-PGA score of 0 or 1 with at least a 2-point improvement from baseline among subjects with a baseline pp-PGA score ≥3~~	
4.1.2.2 Key Secondary Endpoints for Comparisons to Apremilast	The key secondary endpoints for BMS-986165 compared to apremilast at Week 16 are defined as: • sPGA 0/1 response assessed as a proportion of subjects with an sPGA score of 0 or 1 with at least a 2-point improvement from baseline • PASI 75 response assessed as a proportion of subjects who achieve a 75% improvement from baseline in the PASI score • PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in the PASI score • Change from baseline in PSSD symptom score • ss-PGA 0/1 assessed as a proportion of subjects with a ss-PGA score of 0 or 1 with at least a 2-point improvement from baseline among subjects with a baseline ss-PGA score ≥3 • ~~PASI 100 response assessed as a proportion of subjects who achieve a 100% improvement from baseline in the PASI score~~ • sPGA 0 response assessed as a proportion of subjects with an sPGA score of 0 • PSSD symptom score of 0 assessed as a proportion of subjects with a PSSD symptom score of 0 among subjects with a baseline PSSD symptom score ≥1 **The key secondary endpoints for BMS-986165 compared to apremilast at Week 24 are defined as:** • **sPGA 0/1 response assessed as a proportion of subjects with an sPGA score of 0 or 1 with at least a 2-point improvement from baseline at Week 24** • **PASI 75 response assessed as a proportion of subjects who achieve a 75% improvement from baseline in the PASI score at Week 24** • **PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in the PASI score at Week 24** The key secondary endpoints for BMS-986165 compared to apremilast through Week 52 are defined as: • sPGA 0/1 response assessed as a proportion of subjects with an sPGA score of 0 or 1 with at least a 2-point improvement from baseline at Week 52 and **at** Week 24	Updated key secondary endpoints compared to apremilast and re-ordered. Added some clarifying language.

	• PASI 75 response assessed as a proportion of subjects who achieve a 75% improvement from baseline in the PASI score at Week 52 and **at** Week 24 • PASI 90 response assessed as a proportion of subjects who achieve a 90% improvement from baseline in the PASI score at Week 52 and **at** Week 24	
[redacted]		
4.2 Safety	• Treatment-emergent adverse events (TEAEs) – defined as: o ~~New nonserious~~ AEs which ~~first~~ occur after the first dose of study treatment through **30 days** ~~4 weeks~~ after the final dose of the study treatment or subject's participation in the study if the last scheduled visit occurs at a later time.~~; New serious adverse events (SAEs) which first occur after the first dose of study treatment through 4 weeks after the final dose of the study treatment or subject's participation in the study if the last scheduled visit occurs at a later time;~~ o ~~SAEs reported prior to first dose of study treatment that increase in severity or frequency after first dose of study treatment through 4 weeks after the final dose of the study treatment or subject's participation~~	Removed redundant language for treatment-emergent adverse events. Clarified final list for adverse events of interest. Specified that deaths would be evaluated.

	~~in the study if the last scheduled visit occurs at a later time.~~ ○ Treatment-emergent adverse events of interest (AEIs) ~~as determined through the Clinical Safety Program (CSP)~~ for the following events: ▪ Skin-related AEs ▪ Infection AEs**, including influenza** ▪ Creatine kinase (CK) elevation **(evaluated as lab toxicity grade 2 or higher)** ▪ Malignancy ▪ **Deaths** ○ Laboratory abnormalities (as determined by Common Terminology Criteria for Adverse Events [CTCAE **v5.0**] grading) **presented as the worst postbaseline toxicity group** ○ Shifts from baseline **to maximum postbaseline value** ○ **ALT or AST elevation >5 times ULN** ≥ 500 **msec** **Males: < 450 msec, ≥ 450 msec** **Females: < 470 msec, ≥ 470 msec** **In cases where the QT interval is corrected using Bazett's (QTcB) formula is captured instead of QTcF, then QTcB will be converted to QTcF for analyses.** • Eight-Item Patient Health Questionnaire (PHQ-8) total score ○ Absolute and change from baseline values ○ Shifts from baseline scores ○ ~~PHQ-8 total scores ≥ 15~~ ○ Electronic Columbia-Suicide Severity Rating Scale (eC-SSRS) ~~○~~ ~~eC-SSRS items~~**Suicidal ideation and suicidal behavior responses by visit** ~~○~~ **Shifts from baseline to postbaseline value for suicidal ideation and suicidal behavior** ~~○~~ **Worst postbaseline value for suicidal ideation and behavior through Week 16**	Version number added for CTCAE. Provided additional clarifications for lab summaries. Added missing msec for ECG category, new categories for ECG by males and females, and clarification that QTcB will be converted to QTcF for analysis. Removed PHQ-8 total scores >=15 as this is provided in the shift table. Clarified the variables for eC-SSRS to be summarized.
5.0 Populations for Analysis	Full Analysis Set (FAS): All **subjects who are** randomized~~ subjects who are assigned study treatment~~.	The definition for the Full Analysis Set

	Per Protocol Set (PPS): A subset of the FAS who are compliant with study treatment and who do not have any ~~statistically~~ relevant protocol deviations that may impact the coprimary efficacy endpoint assessments.	was modified to clarify that the population includes all randomized subjects.Removed "statistically" from relevant deviations in the PPS description to align with standard naming.
5.1 Relevant Protocol Deviations	Relevant protocol deviations are deviations that can have an impact on the primary efficacy endpoints. The impact of relevant protocol deviations on the primary efficacy results will be assessed by excluding subjects from the FAS to define the PPS in supportive analyses of the primary efficacy endpoints. Relevant protocol deviations to be considered regarding exclusion of subjects from the FAS will include the following: • Randomized but did not take any study treatment • **No postbaseline PASI or sPGA** • **Baseline BSA involvement < 10%** • **Baseline PASI score < 12** • **Baseline sPGA < 3** • **Did not have non-plaque psoriasis at baseline** • ~~Failed to meet study inclusion criteria but were entered into the study:~~ • ~~Met study exclusion criteria but were entered into the study: (only exclusion criteria expected to have an impact on the primary efficacy endpoints will be considered relevant)~~ • Poor compliance to study medication within the first 16 weeks of treatment defined as <~~80~~**75**% compliant with study • Failure to adhere to prohibited concomitant medication restrictions **as described below**: ○ **Investigational drug or placebo taken outside of the study any time between Day 1 and the Week 16 assessment** ○ **Phototherapy within 4 weeks prior to the Week 16 assessment** ○ **Biologic medications (eg, adalimumab, etanercept, infliximab, ustekinumab) any time between Day 1 and the Week 16 assessment** ○ **Oral psoriasis medications any time between Day 1 and the Week 16 assessment** ○ **Oral corticosteroids (unless for the treatment of an adverse event) within 4 weeks prior to the Week 16 assessment**	Deviations were updated with final categories.

	○ **Topical medications/treatments that could affect psoriasis evaluations within 2 weeks prior to the Week 16 assessment** ○ **Medicated shampoos within 2 weeks prior to the Week 16 assessment** • ~~Subject overdosed, misused or abused study treatment prior to Week 16~~ • ~~Actual treatment received is different than randomized treatment~~**Subject received treatment that was different than intended treatment at any visit prior to Week 16.** All subjects with relevant protocol deviations will be identified prior to database lock and unblinding of treatment assignment **with the exception of subjects where actual treatment received is different than randomized treatment which will be determined after treatment unblinding**. Relevant protocol deviations will be summarized by treatment group and deviation category for the FAS population. ~~Additionally, all important protocol deviations, which are deviations that may impact the efficacy and safety of subjects, will be identified prior to database lock and unblinding of treatment assignment. Important protocol deviations will be summarized by treatment group and deviation category for the FAS population.~~	
6.1 Efficacy Analyses	Coprimary and key secondary efficacy endpoints will be displayed graphically by treatment group and visit, as applicable. ~~Binary endpoints will be displayed using bar charts and continuous endpoints will be displayed using line plots.~~ All efficacy endpoints will be summarized descriptively. The key secondary endpoints will be **tested with a** two-sided ~~with a~~ significance level of 0.025 to compare BMS-986165 6mg QD vs. placebo and vs. apremilast ~~(equivalent to 0.05 level of significance)~~. A hierarchical **testing** approach will be used ~~in~~ **for** testing of key secondary endpoints **(see Section 6.1.3)**~~within each comparison branch~~. Two-sided 95% confidence intervals (CIs) will be provided for all efficacy estimates. [redacted] ~~The analyses of the primary and secondary endpoints of all subjects including those from Japan and China will include baseline body weight as a factor in the model. Therefore, although the subjects coming from Japan and China may not be balanced by the body weight groups (<90 kg or ≥90 kg), analyses of the endpoints will be adjusted for any effect of the baseline body weight on efficacy.~~**Due to small sample sizes in China and Japan and differences in stratification factors, analyses will be stratified by a combination of the stratification factors used in randomization. The stratified levels are provided below:** ▪ **US/PRIOR USE/< 90 KG** ▪ **US/PRIOR USE/>=90 KG** ▪ **US/NAÏVE/<90 KG** ▪ **US/NAÏVE/>=90 KG**	Removed plot description to allow flexibility in data presentation plan. Clarified the language for testing of secondary endpoints. Updated the stratification levels to a combination stratification factor due to small sample sizes in China and Japan.

	▪ **ROW/PRIOR USE/< 90 KG** ▪ **ROW/PRIOR USE/>=90 KG** ▪ **ROW/NAÏVE/<90 KG** ▪ **ROW/NAÏVE/>=90 KG** ▪ **JAPAN/PRIOR USE** ▪ **JAPAN/NAÏVE** ▪ **CHINA/PRIOR USE** ▪ **CHINA/NAÏVE** **If there is not at least one subject per treatment group in each of the China and Japan strata, then these regions will be collapsed together to form the following stratified levels for analyses:** ▪ **US/PRIOR USE/< 90 KG** ▪ **US/PRIOR USE/>=90 KG** ▪ **US/NAÏVE/<90 KG** ▪ **US/NAÏVE/>=90 KG** ▪ **ROW/PRIOR USE/< 90 KG** ▪ **ROW/PRIOR USE/>=90 KG** ▪ **ROW/NAÏVE/<90 KG** ▪ **ROW/NAÏVE/>=90 KG** ▪ **ASIA/PRIOR USE** ▪ **ASIA/NAÏVE**	
6.1.1 Primary Endpoints 6.1.1.1 Primary Analysis	Analysis Model A stratified Cochran-Mantel-Haenszel (CMH) test will be used to compare the sPGA 0/1 response rates at Week 16 between BMS-986165 6 mg QD and placebo **using the stratification factors from IRT specified in Section 6.1**. ~~The following prognostic factors are included in the model: geographic region (U.S., Japan, China, Rest of World), body weight (<90 kg, ≥90 kg) and prior biologic use (yes/no).~~ The odds ratio (ratio of odds in BMS-986165 6 mg QD group to the odds in the placebo group) and the corresponding 2-sided 95% CIs and p-values will be reported. Additionally, treatment differences in proportions and the 95% CI for the treatment difference will be provided. The common treatment difference in proportion with the confidence interval, common odds ratio with the confidence interval and p-value will be estimated by the Mantel-Haenszel method consistently. A similar analysis will be performed to compare the PASI 75 response rates at Week 16 between BMS-986165 6mg QD and placebo. Imputation Methodology Non-responder imputation (NRI) will be used for coprimary efficacy endpoints for subjects who: • Discontinue treatment or study prior to Week 16	Clarified stratification to be used and removed prohibited medication criteria.

	~~• Start a protocol prohibited medication/therapy that could improve psoriasis during the first 16 weeks of treatment, or~~ • Have missing Week 16 endpoint data for any reason ~~Subjects taking protocol prohibited medications during the first 16 weeks of treatment will be identified prior to database lock and treatment unblinding.~~ NRI will be used as the primary method of imputation for the coprimary efficacy endpoints.	
6.1.1.2 Sensitivity Analyses	As a method to assess the sensitivity of the primary imputation method for the coprimary endpoints, further imputation methods will be used to impute Week 16 data in subjects who: • Discontinue treatment or study prior to Week 16 ~~• Start a protocol prohibited medication/therapy that could improve psoriasis during the first 16 weeks of treatment, or~~ • Have missing Week 16 endpoint data for any reason <u>Tipping Point Analysis</u> Tipping point analysis will further be used to assess the robustness of the primary study results. Different number of events between treatment groups will be assessed until the study conclusion is changed. Each imputed value is initially imputed as a responder. The imputed values in the placebo group will remain as responders while the BMS-986165 imputed values are replaced as a nonresponders one at a time, therefore changing the number of events between groups. Once all imputed values in the BMS-986165 group have been replaced with nonresponders values, the data will reset to where the BMS-986165 group imputed values are all responders and one by one the imputed values in the placebo group are replaced with a nonresponder until all placebo are nonresponders and all BMS-986165 are responders. Furthermore, every pair of imputations between the placebo and BMS-986165 groups will be assessed similarly, creating a matrix of possible patterns. At each iteration, the statistical analysis is assessed, and the direction of the analysis is recorded. A graph will be provided to identify at what point to which the difference in events cause a direction shift in study results. **The tipping point analyses will be based on a two-sample chi-square test approach rather than a stratified CMH test approach for simplicity of the analysis.** **<u>Multiple Imputation</u>** **Multiple imputation (MI) will be used for sensitivity analyses for each of the coprimary efficacy endpoints, PASI 75 response at Week 16 and sPGA 0/1 response at Week 16. Multiple imputation of missing data for PASI 75 response and sPGA 0/1 response will be performed by fully conditional specification (FCS) using PROC MI in SAS with the FCS option. The FCS MI specifies the multivariate imputation model on a variable-by-variable basis by a set of conditional densities. For each missing value at Week *X*, the FCS MI will generate values from a conditional distribution for the missing data given the other data prior to**	Removed prohibited medication criteria. Provided some clarification for the tipping point analysis. Added a sensitivity analysis for multiple imputation.

	Week *X*. The FCS MI model will include treatment group and stratification factors for randomization. A total of 1000 imputed (complete) datasets will be generated for the MI analysis. The same test procedure used for the main analysis (CMH test), will be used to analyze the responder endpoint from each imputed dataset. SAS PROC MIANALYZE will be used to pool the results from the CMH tests and generate an overall result by Rubin's rules (1987).	
6.1.1.3 Supportive Analyses	~~Additional Analysis~~ ~~Additionally, a logistic regression model will be used to compare the sPGA 0/1 response rates at Week 16 between BMS-986165 6mg QD and placebo. The prognostic factors will be included in the model as covariates. The estimated odds of response with corresponding 2-sided 95% CIs and p-values will be reported. A similar analysis will be performed on the PASI 75 response rates at Week 16. These analyses will be performed on the FAS.~~	Removed this analysis since it similar to CMH analysis.
6.1.2 Key Secondary Endpoints 6.1.2.1 Binary Endpoints	Analysis Model CMH tests will be used to compare response rates between either BMS-986165 6mg QD and placebo or apremilast **using the stratification factors from IRT specified in Section 6.1**. ~~The following prognostic factors are included in the model: geographic region (U.S., Japan, China, Rest of World), body weight (<90 kg, ≥90 kg) and prior biologic use (yes/no).~~ The odds ratio (ratio of odds in BMS-986165 6 mg QD group to the odds in the placebo **or apremilast** group) and the corresponding 2-sided 95% CIs and p-values will be reported. Additionally, treatment differences in proportions and the 95% CI for the treatment difference will be provided ~~from Chi-square tests~~. Imputation Methodology for Week 16 Endpoints NRI will be used for ~~coprimary~~**key secondary** efficacy endpoints for subjects who: • Discontinue treatment or study prior to Week 16 • ~~Start a protocol prohibited medication/therapy that could improve psoriasis during the first 16 weeks of treatment, or~~ • Have missing Week 16 endpoint data for any reason ~~Subjects taking protocol prohibited medications during the first 16 weeks of treatment will be identified prior to database lock and treatment unblinding. NRI will be used as the primary method of imputation for the coprimaryefficacy endpoints.~~ **Imputation Methodology for Week 24 Endpoints (for comparison of BMS-986165 6 mg QD to apremilast)** **NRI will be used for key secondary efficacy endpoints for subjects who:** • **Discontinue treatment or study prior to Week 24** • **Have missing Week 24 endpoint data for any reason**	Clarified stratification to be used and removed prohibited medication criteria. Corrected coprimary to key secondary. Added imputation for Week 24 endpoints. Add clarifications for imputations of Week 52 endpoints and definition for meeting response criteria.

	Imputation Methodology for Week 52 Endpoints **(for comparison of BMS-986165 6 mg QD to apremilast)** **The key secondary endpoints included in this analysis are the following:** • **sPGA 0/1 response with at least a 2-point improvement from baseline at Week 52 and at Week 24** • **PASI 75 response at Week 52 and at Week 24** • **PASI 90 response at Week 52 and at Week 24** **Subjects must meet response criteria at both Weeks 24 and 52 in order to be considered a responder. Those who do not meet response criteria at both weeks will be considered a nonresponder.** NRI will be used for **the above** key secondary endpoints for subjects who: • Discontinue treatment or study prior to Week 52 • ~~Start a protocol prohibited medication/therapy that could improve psoriasis during the first 52 weeks of treatment, or~~ • Have missing Week 52 or Week 24 endpoint data • **Switch from apremilast to BMS-986165 at Week 24 for Week 24 PASI 50 nonresponders** ~~Subjects taking protocol prohibited medications will be identified prior to database lock and treatment unblinding.~~	
6.1.2.2 Continuous Endpoints	Analysis Model Analysis of covariance (ANCOVA) models will be used to compare the change between BMS-986165 6 mg QD and placebo or between BMS-986165 6 mg QD and apremilast **at Week 16**. Treatment group will be included in the model as well as the **stratification factors from IRT specified in Section 6.1**~~following baseline prognostic factors: geographic region (U.S., Japan, China, Rest of World), body weight (<90 kg, ≥90 kg) and prior biologic use (yes/no)~~ and will be considered as fixed effects. The baseline value will be added into the model as a covariate. The adjusted least-squares (LS) means as well as the treatment differences based on LS means and the corresponding 2-sided 95% CIs will be provided for the difference between BMS-986165 6 mg QD and placebo or apremilast. Imputation Methodology **for Week 16 Endpoints** For continuous **key** secondary endpoints, a modified baseline observation carried forward (mBOCF) approach will be used for missing data. The baseline observation will be carried forward for subjects who discontinue study treatment **prior to Week 16** due to: • Lack of efficacy • AEs Subjects who discontinue study treatment **prior to Week 16** for other reasons **or who have a missing Week 16 value** will have the last valid observation carried forward (including the baseline value as applicable).	Clarified stratification to be used and removed prohibited medication criteria. Added a clarification that subjects with missing baseline values will be excluded from the analysis for change from baseline endpoint.

	~~For subjects who start a protocol prohibited medication/therapy that could improve psoriasis prior to the endpoint will also have their endpoint value imputed as the baseline value. The last valid observation will be carried forward for all other subjects with missing data.~~ **Subjects with a missing baseline value will be excluded from the analysis for the change from baseline endpoint.**	

[redacted]

[redacted]

[redacted]

6.1.4 Subgroup Analyses	Subgroup analyses will be performed on the coprimary endpoints using the FAS. The primary imputation method will be applied for these analyses. The CMH test The CMH test **using the stratification factors from IRT** will be the analysis method used where the stratification factor is the specified subgroup.will be the analysis method used ~~where the stratification factor is the specified subgroup~~. The following subgroups will be considered: • Geographic region (U.S., Japan, China, Rest of World) • Country • Sex (male, female) • Age group (<65 y, ≥65 y) • Body weight (<90 kg, ≥90 kg) **– from case report form** • Race • Prior biologic use for psoriasis, psoriatic arthritis and other inflammatory diseases only (yes, no) **– from case report form** • **Prior systemic treatment use (yes/no)** • **Prior phototherapy use (yes/no)** • sPGA (3, 4) • PASI score (≤20, >20) • BSA involvement (10-20, >20)	Removed some subgroups that will not be analyzed. Clarified weight and prior biologic use to be taken from CRF and stratification factors for the CMH statement will use IRT stratification factors.

	~~• DLQI (0-5, 6-10, 11-30)~~ • Duration of disease (< 10 y, ≥ 10 y) • Age at disease onset (<18, 18-39, ≥40) ~~• Smoking status:~~ ~~Currently non-smoker – never smoked or quit smoking ≥6 months from the 1st treatment~~ ~~Currently non-daily smoker or light smoker - smoked on no more than 25 days in the previous 30 days to < 5 cigarettes per day~~ ~~• Currently moderate to heavy smoker – ≥10 cigarettes per day~~	

[redacted]

[redacted]

[redacted]

[redacted]

6.2.1 Adverse Events	Adverse events ~~(AE)~~ will be presented for the number and percentage of subjects and the number of events. ~~All adverse events (t~~**T**reatment-emergent ~~[TEAE] and non-treatment emergent)~~ will be provided in listings. Summary tables will be reported in decreasing frequency based on the ~~total~~ **BMS-986165** column. Counting for frequency analysis will be by subject and not by AEs, and subjects are only counted once for recurring AEs within each system organ class (SOC) or preferred term (PT); however, the number of total AEs per subject, including multiple occurrences of individual AEs, will also be presented. For summaries of severity, AEs will be reported only at their maximum severity in each subject. Subjects with multiple AEs within a particular SOC or PT will be counted under the category of their most severe event within that SOC or PT. An overall summary for the following categories will be presented: • ~~Subjects with at least one TEAE~~ • ~~Subjects with at least one related TEAE~~ • ~~Subjects with at least one treatment-emergent SAE~~ • ~~Subjects with at least one related treatment-emergent SAE~~ • ~~Subjects discontinuing study treatment due to a TEAE~~ • ~~Subjects discontinuing from study due to a TEAE~~ • ~~Subjects who died due to an AE~~ • ~~Subjects who died due to a TEAE~~ • **Deaths** • **SAEs** • **Related SAEs** • **AEs** • **Related AEs** • **Discontinued treatment due to AEs** ~~A summary of TEAEs leading to discontinuation of study treatment or study will be provided, grouped by SOC and PT for all TEAEs and treatment-related TEAEs.~~	Modified text for overall summary to align with the data presentation plan. TEAEs leading to discontinuation was removed from this section as this is described in a separate section.
6.2.1.1 Adverse Events of Interest (AEI) **and Other Important AEs**	Summaries for treatment-emergent AEIs events will be provided **by PT** for each AEI category~~for the following~~: • ~~TEAEs by SOC and PT~~ ~~Skin Events~~ ~~The number and percentage of subjects reporting each type of skin event and the corresponding location will be summarized.~~ ~~Infections~~	Modified the text for the final list of adverse events of interest and the adjudicated cardiovascular events.

	~~The number and percentage of subjects reporting infections will be summarized.~~ ~~Creatine kinase (CK) elevation~~ ~~The number and percentage of subjects reporting CK elevations will be summarized.~~ ~~Malignancies~~ ~~The number and percentage of subjects reporting malignancies will be summarized.~~ ~~Cardiovascular~~ ~~The number and percentage of subjects reporting cardiovascular event will be summarized.~~ • **Skin-related events** • **Infection events** • **Malignancy events** **Creatine kinase (CK) elevation for CK elevation >2.5 upper limit of normal will be summarized as CTCAE grade 2 or higher in the clinical laboratory summaries.** **Additional information collected for some events as part of the clinical safety program and adjudicated events will also be summarized.**	
6.2.1.2 Serious Adverse Events	Summaries for treatment-emergent SAEs will be provided for the following: • Treatment-emergent SAEs by SOC and PT • ~~Treatment-related treatment-emergent SAEs by SOC and PT~~	Aligned with the Data Presentation Plan
6.2.1.3 Adverse Events Leading to Discontinuation of Study Treatment **or Study Treatment Interruption**	Summaries for TEAEs leading to discontinuation of study treatment will be provided for the following: • TEAEs by SOC and PT • ~~Treatment-related TEAEs by SOC and PT~~ **Summaries for TEAEs leading to study treatment interruption will be provided for the following:** • **TEAEs by SOC and PT**	Added summary for TEAEs leading to study treatment interruption and aligned with the Data Presentation Plan.
6.2.3 Clinical Laboratory Data	Laboratory parameters will be summarized using the International System (SI) of Units~~, unless otherwise specified~~ **and US conventional units**. Data will be summarized by time point, as applicable. The following summaries will be provided for each parameter: • Absolute and change from baseline values for continuous parameters • Number and percentage of subjects for the following: ~~○ Categorical urinalysis parameter results~~	Added summaries for US conventional units. Removed urinalysis categorical summaries to align with data

	○ Maximum postbaseline CTCAE grade for each applicable laboratory parameter **through Week 16** ○ Shifts from baseline based on maximum postbaseline CTCAE grade **through Week 16** • Drug-induced Liver Injury (DILI) and Hy's Law summaries All laboratory data specified in the summary tables will be present in listings.	presentation plan. Added additional clarifications to align with data presentation plan.
6.2.4 Vital Signs and Physical Findings	Vital signs**, including weight,** will be summarized by time point, as applicable. The following summaries will be provided for each parameter:	Clarified that weight will be summarized.
6.2.6 Other Safety Data 6.2.6.1 PHQ-8	PHQ-8 total score will be summarized by time point, as applicable. The following summaries will be provided: • Absolute and change from baseline values • Number and percentage of subjects: ○ Shifts from none, mild, moderate, moderately severe and severe scores at baseline and at each time point ~~○ PHQ-8 total scores ≥ 15~~	Removed PHQ-8 total score >=15 summaries as this is provided in the shift summary.
6.2.6.2 eC-SSRS	Suicidal ideation and behavior individual item responses will be summarized by time point, as applicable. The following summaries will be provided: • Number and percentage of subjects with positive responses on suicidal ideation and/or suicidal behavior **questions**~~responses~~ **for each question and overall all questions within suicidal ideation and suicidal behavior** • **Shifts from baseline based on maximum postbaseline response through Week 16** • **Worst postbaseline value for suicidal ideation and behavior through Week 16**	Additional analyses have been added for eC-SSRS.
6.3.1 Subject Populations and Disposition	The number of subjects enrolled/screened and the number and percentage of subjects randomized, treated, and in each analysis population will be presented. The number and percentage of subjects randomized in each region, country, and site will be presented. Additionally, the following summaries will be provided for the FAS by treatment group and overall: ~~• Number and percentage of subjects on treatment at each visit week~~	Aligned this section with the summaries in the Data Presentation Plan
6.3.2 Demographic and Baseline Characteristics	Demographic **and baseline** characteristics will be summarized by treatment group for **the FAS** ~~all of the analysis populations~~**.** Demographic and baseline characteristics include the following: • **Prior systemic treatment use (yes/no)** • **Prior phototherapy use (yes/no)**	Updated populations to be used for summaries and removed baseline

	~~• DLQI (0-5, 6-10, 11-30)~~ ~~• Smoking status (Current daily smoker and whether heavy vs. light, Current occasional smoker, former smoker, never smoker, smoker current status unknown, unknown if ever smoked)~~ Additional demographics or baseline data may be added to summary tables. General medical history and medical history related to PsA will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). General medical history data will be summarized for each SOC and PT by treatment group and overall for the FAS ~~and Safety Set populations~~. Separate tables will be provided for ~~general~~ **psoriasis** medical history.	characteristic that will not be summarized.
6.3.3 Prior and Concomitant Medications	Prior and concomitant medications will be coded according to World Health Organization-Drug Dictionary WHO-DD (version at the time of DBL) and will be summarized by Anatomic Therapeutic Classification (ATC) and preferred term (PT) by treatment group for the **As-treated population**~~FAS~~. The number and percentage of subjects using at least one medication and each medication will be displayed by treatment group. Subjects taking more than one medication within the same ATC or PT will be counted once.**Summaries will be provided for prior medications and medications that started prior to first treatment and were ongoing after first treatment start date as well as concomitant medications.**	Clarified summaries to be provided and the population to be used for the summaries.
6.3.3.1 Psoriasis, Psoriatic Arthritis and Other Inflammatory Disease Related Systemic Medications	Prior **medications and medications that started prior to first treatment and were ongoing after first treatment start date for** systemic biologic and non-biologic medications will be summarized as described above for the number and percentage of subjects using each reported medication. Concomitant use of DMARDs like methotrexate will be summarized.	Added clarification for prior and ongoing medications to be summarized.
6.3.3.2 Concomitant Corticosteroid Use	Additionally, corticosteroids will be summarized by ATC and PT ~~for the FAS~~.	Removed reference to analysis population as this is redundant information with Section 6.3.3.
6.3.4 Exposure 6.3.4.1 Duration of Treatment	Duration by GroupOverall duration of each treatment, in days, will be calculated for each subject in each treatment group. Subjects are dispensed study treatment at each visit, starting with Day 1 (Week 0) and at each subsequent visit~~. Subjects are also to record daily dosing in an eDiary.~~ The ~~Day 1 dispensed date will be considered as the~~ date of first dose of study treatment **is the Week 0 PK dosing date** and is recorded on the eCRF. **If this date is missing, then the**	Updates to duration of exposure formula's were made to align with the manner in

	earliest drug dispensation date will be used. The last **dose** date ~~of dose~~ is ~~defined as the last day a subject received drug and is~~ recorded on the end of treatment page in the eCRF. **If this date is missing, then the latest date from the [redacted] page or drug accountability return date will be used.** ~~If the date of last dose of study treatment is missing, i.e.: subject lost to follow-up, the end of study date or last contact date (whichever comes first) will be used to calculate duration of treatment.~~ Duration of treatment will be summarized descriptively by **randomized** treatment group. Placebo:For subjects randomized to placebo, ~~the Week 16 date will be used as the date of last dose of placebo. Formula for~~ duration is defined as: • Placebo = **Date of last dose of placebo** ~~Week 16 date~~ – date of first dose +1 • BMS-986165 = Date of last dose of BMS-986165 – ~~Week 16 date~~**date of first dose of BMS-986165 +1** ~~If a placebo subject discontinues study treatment on or before the Week 16 visit, the date of last dose will be used to calculate the duration of placebo.~~ **Apremilast:** For subjects randomized to apremilast, the Week 24 date will be used as the date of last dose of apremilast and the date of first dose of either apremilast or BMS-986165 for the subsequent period. Subjects who receive apremilast at Week 24 through to Week 52 will be counted as a sum of the first 24 weeks of treatment and the next treatment period. • Apremilast (Wk 0 through Wk 24) = Week 24 date – date of first dose +1 • Apremilast (Wk 24 through Wk 52) = Date of last dose – Week 24 date • Total Apremilast = **Duration of ~~A~~apremilast** Wk 0-24 + **duration of** apremilast Wk 24-52 BMS-986165 = Date of last dose of BMS-986165 – Week 24 date Duration by Period Overall duration (in days) of each treatment received within each study period will be calculated. Duration within each period is defined for each treatment group as: $$Last\ dose\ date\ \boldsymbol{in\ period} - first\ dose\ date\ \boldsymbol{in\ period} + 1$$ ~~Subjects are dispensed study treatment at each visit, starting with Day 1 (Week 0) and at each subsequent visit. Subjects are also to record daily dosing in an eDiary. The dispensed drug date of Day 1 of the study period will be considered as the date of first dose of study treatment for that period and is recorded on the eCRF. The last date of dose within the period will be considered as the day prior to the next period start date.~~	which data will be summarized.

	~~If the date of last dose of study treatment is missing, i.e.: subject lost to follow-up, the end of study date or last contact date (whichever comes first) will be used to calculate duration of treatment.~~	
6.3.4.2 Summary of Dosing	The number of doses taken for each subject for each period will be determined using the eCRF drug accountability data and is defined as: $Doses\ Taken = \frac{(number\ of\ tablets\ dispensed - number\ of\ tablets\ returned)}{~~3~~}$ The number of doses taken will be summarized descriptively by treatment group **within each study period and overall**.	Clarified how data will be summarized.
6.3.4.3 Compliance	Treatment compliance will be determined from data captured on the Drug Accountability eCRF. ~~The number and percentage of subjects who have missed at least one dose will be provided.~~ ~~Additionally, descriptive statistics for the number of missed doses within each treatment period and overall will be provided by treatment group. The number of missed doses for each subject will be calculated for each period.~~ $Number\ of\ expected\ doses: (date\ of\ next\ visit - date\ of\ current\ visit) x3$ ~~$Number\ of\ missed\ doses: Number\ of\ expected\ doses - number\ of\ doses\ taken$~~ Treatment compliance will be derived for each period ~~and overall~~. Compliance is defined as: $\left(\frac{Number\ of\ doses\ taken}{Number\ of\ expected\ doses}\right) x100$ Period compliance will be calculated by summing over all visits within the period ~~and overall compliance will be calculated by summing each visit in the study~~ and summarized using descriptive statistics by treatment group. The number and percentage of subjects with <~~80~~**75**%, ~~80~~**75**% to 1**20**0%, and >1**20**0% compliance will be provided by treatment group for each ~~visit,~~ period ~~and overall~~. **If a subject does not return the container, then this dispensation event will be excluded in the calculation for the total number of doses taken, total number of expected doses, and total compliance.**	Clarified how data will be summarized.

[redacted]

[redacted]

[redacted]

[redacted]

6.6 Statistical Impacts Due to COVID-19 6.61 Impact on Efficacy Endpoints 6.6.2 Impact on Safety Endpoints	**Statistical Impacts Due to COVID-19** **Impact on Efficacy Endpoints** **There are no COVID-19-related impacts to the Week 16 and Week 24 endpoints as all subjects remaining in the trial completed these visits prior to COVID-19 restrictions. Key secondary efficacy endpoints involving later visits may be impacted by COVID-19 in the form of missing data and/or a treatment assignment defaulting to BMS-986165. If a subject has missing data or is switched to BMS-986165 from a different treatment due to Covid-19 during the Week 24 through Week 52 period, the subject will be excluded from the analysis.** **Key secondary efficacy endpoints that may be impacted include the following:** • **sPGA 0/1 response at Week 52 and at Week 24** • **PASI 75 response at Week 52 and at Week 24** • **PASI 90 response at Week 52 and at Week 24** **A sensitivity analysis will also be performed for the above endpoints using the last observed value. The data handling rules to be used are as follows for subjects with missing data or for those who switch to BMS-986165 from a different treatment due to COVID-19:** • **If the subject is missing the Week 52 response value (ie, sPGA 0/1, PASI 75, PASI 90), then the last observed response value during the Week 24 through Week 52 period will be carried forward to the Week 52 value** • **If the subject is switched to BMS-986165 from a different treatment due to COVID-19, then the last observed response value prior to the switch will be carried forward to the Week 52 value** [redacted]	New section added to address COVID-19.

	[redacted] **Time to first loss endpoints include:** • **Time to first loss of PASI 75 among subjects that are PASI 75 responders at Week 24** • **Time to first loss of sPGA 0/1 among subjects that are sPGA 0/1 responders at Week 24 where loss of sPGA 0/1 is defined as an sPGA score ≥ 2 (sPGA scores are rounded to the nearest whole number)** **Impact on Safety Endpoints** **No modifications will be made for the analysis of safety data due to COVID-19-related issues. There will not be any additional safety data handling considerations for missing data or treatment interruptions due to COVID-19.**	
7.2 Sequence of Planned Analyses	[redacted]	
7.2.2 Final Analyses and Reporting	**Analyses of the palmoplantar pp-PGA 0/1 endpoint are provided for descriptive purposes within the individual CSR due to small sample sizes. A pooled analysis of data from IM011046 and IM011047 will be conducted for the submission.**	Added clarification for the analyses of palmoplantar (pp-PGA 0/1 endpoint).
8.1 General Definitions	Study day is calculated as: assessment date – date **of first dose** ~~the subject is randomized~~ + 1 Baseline - Unless otherwise stated, Baseline is defined as the ~~last~~ measurement ~~prior to dosing on Day 1~~ **at the randomization visit** (Week 0). If the measurement **at the randomization visit**~~on Day 1~~ is missing~~ or not available~~, then a prior measurement during the screening period may be used as baseline. Baseline assessments must be performed per protocol and standard of care assessments may not be used for baseline. **Baseline for each PSSD question will be calculated as the average value over the 7 days prior to the randomization visit. In the event of missing values in the 7 days prior to the randomization visit, daily scores from at least 4 days out of the 7 days can be used to calculate the average score. If > 3 daily scores in the last 7 days are missing for a particular question, then the baseline value for that question will be missing. If any questions are missing baseline after using the described rule, then the baseline value for the subscore using that**	Updated first and last dose date definitions and other definitions. Added an additional rule for percent change from baseline.

	question and the baseline value for the total score will be missing.
	
	First Dose Date – Study: The date a subject received their first dose on Day 1 as recorded in the eCRF ~~as date study treatment was dispensed~~**Week 0 [redacted] dosing date or the earliest drug dispensation date**.
	
	First Dose Date – Period: The date a subject received their first dose ~~in the defined study period~~ as recorded in the eCRF ~~as date study treatment was dispensed~~**Week 0 [redacted] dosing date or the earliest drug dispensation date for Treatment Period 1 and earliest drug dispensation date for Treatment Periods 2 and 3**.
	
	Change from baseline in the maximum post-baseline value is defined as highest observed value **or grade** post-baseline. The change is calculated using this value as the post-baseline value.
	
	Change in the maximum post-baseline value **or change in the worst post-baseline value: Change from baseline in the worst post-baseline value is defined as worst observed value post-baseline. The change is calculated using this value as the post-baseline value.**
	
	If the baseline value is 0 and the post-baseline is >0, then the percent change from baseline value will be missing.
	
	EAIR:
	
	EAIR = 100*~~The number of subjects with a specific event divided by the total exposure time (in years) among the subjects in the treatment group, where the total exposure time in years is calculated by dividing the sum of exposure time in days over all subjects included in the treatment group*~~365.25***(total number of subjects with the AE)/total exposure time for the selected AE under each treatment that a subject is exposed. Where total exposure time for each AE within a treatment is calculated as follows:**
	
	• **If a subject has at least 1 event while on a particular treatment, then the exposure time for that subject and AE on that treatment is:**
	  ○ **First AE onset date – treatment start date (of that particular treatment) +1**
	• **If a subject does not have an event, exposure time for that AE is:**
	  ○ **Treatment stop date (of that particular treatment) – Treatment start date (of that particular treatment) + 1 + 30 days (if subject discontinued or subject completed Period 3 and is not rolling into IM011075 study)**
	  ○ **Treatment stop date (of that particular treatment) – Treatment start date (of that particular treatment) + 1 (if subject completed Period 3 and is rolling into IM011075 study)**
	• **Total exposure time = sum of exposure time for each AE within a treatment**	

8.2.1.3 Body Surface Area (BSA)	The product of BSA and sPGA will be **calculated**~~derived as a potential proxy measure to PASI scores~~.	Updated the product score language.
8.2.2.1 Psoriasis Symptoms and Signs Diary (PSSD)	**Baseline for each PSSD question will be calculated as the average value over the 7 days prior to the randomization visit. In the event of missing values in the 7 days prior to the randomization visit, daily scores from at least 4 days out of the 7 days can be used to calculate the average score. If > 3 daily scores in the last 7 days are missing for a particular question, then the baseline value for that question will be missing. If any questions are missing baseline after using the described rule, then the baseline value for the subscore using that question and the baseline value for the total score will be missing.**	Clarified baseline definition for PSSD.

[redacted]

[redacted]

8.2.2.11 Psoriatic Arthritis Screening and Evaluation (PASE) Questionnaire	The PASE questionnaire will be administered at Screening in subjects with peripheral joint complaints. The PASE questionnaire is a self-administered tool that is used to screen for psoriatic arthritis among subjects who have psoriasis. The PASE questionnaire consists of 15 questions subdivided into 7 questions focusing on symptoms of psoriatic arthritis, and 8 questions focusing on the impact of psoriatic arthritis on function. Each question is scored on a 1 to 5 scale, with a maximum symptom sub-scale score of 35, a maximum function sub-scale score of 40 giving a total maximum score of 75. A score of 47 or above has been shown to distinguish between psoriatic arthritis and non-psoriatic arthritis. This questionnaire is only intended to be a screening tool for psoriatic arthritis and does not replace a comprehensive musculoskeletal exam performed by a rheumatologist. **This information will not be summarized.** ~~Individual score to each of the 15 questions will be collected in the eCOA system. The sub-scale scores and total score will be derived in the analysis datasets.~~	Clarified that PASE information will not be summarized.
8.2.2.13 electronic Columbia-Suicide Severity Rating Scale (eC-SSRS)	The electronic Columbia-Suicide Severity Rating Scale (eC-SSRS) is a computer-automated, patient-reported version of the C-SSRS instrument that defines 11 categories of suicidal ideation and behavior (SIB) events. Categories and definitions are provided in Appendix 23 of the protocol. The categories are as follows: • Suicidal ideation 6. **Wish to be dead**~~Passive~~ 7. **Non-specific active suicidal thoughts**~~Active: Nonspecific (no method, intent, or plan)~~ 8. **Active suicidal ideation with any methods without intent to act**~~Active: Method, but no intent or plan~~	Updated descriptions of categories to align with data presentation plan and put behavior categories in proper ascending order.

	9. **Active suicidal ideation with some intent to act, without specific plan**~~Active: Method and intent, but no plan~~ 10. **Active suicidal ideation with specific plan and intent**~~Active: Method, intent, and plan~~ • Suicidal behavior 6. **Preparatory acts or behavior**~~Completed suicide~~ 7. **Aborted attempt**~~Suicide attempt~~ 8. Interrupted attempt 9. **Actual attempt (Non-fatal)**~~Aborted attempt~~ 10. **Completed suicide**~~Preparatory actions toward imminent suicidal behaviors~~ Self-injurious behavior, no suicidal intent	
8.4 Study Periods	Period 1 = ~~the first 16 weeks of treatment~~**Week 0 to Week 16 visit date** Period 2 = **Week 16 visit date +1 to Week 24 visit date** ~~after Week 16 to Week 24~~ Period 3 = **Week 24 visit date +1 to Week 52 visit date**~~after Week 24~~ **Follow-up = 4 week follow-up period**	Clarified study periods.
8.5 Day Ranges for Analysis Visits	Week 16 100, 127 (or Week 16 drug dispense date ~~if earlier~~) Week 24 156, 183 (or Week 24 drug dispense date ~~if earlier~~)	Clarified Week 16 and Week 24 analysis range criteria

Appendix 1 Planned Analyses

List of Planned Analyses: Comparisons of BMS-986165 6 mg QD vs. Placebo		
Measure of Interest	**Population**	**Analysis at Week 16**
sPGA 0/1 – coprimary PASI 75 - coprimary	FAS	NRI+CMH - primary
sPGA 0/1 – coprimary PASI 75 – coprimary	FAS	LOCF+CMH – sensitivity LOCF/NRI+CMH – sensitivity Tipping Point+CMH – sensitivity Multiple Imputation + CMH - sensitivity
sPGA 0/1 – coprimary PASI 75 – coprimary	PPS	NRI+CMH - supportive
sPGA 0/1 – coprimary PASI 75 - coprimary	FAS	NRI+CMH – subgroups
PASI 90 – key secondary	FAS	NRI+CMH
sPGA 0 – key secondary	FAS	NRI+CMH
ss-PGA 0/1 among subjects with a baseline ss-PGA ≥ 3 – key secondary	FAS	NRI+CMH
PSSD symptom score of 0 – key secondary	FAS	NRI+CMH
DLQI 0/1 – key secondary (EX-US submission only)	FAS	NRI+CMH
PGA-F 0/1 among subjects with a baseline PGA-F ≥ 3	FAS	NRI+CMH

List of Planned Analyses: Comparisons of BMS-986165 6 mg QD vs. Apremilast			[redacted]
Measure of Interest	**Population**	**Primary Analysis at Week 16/24/52**	
sPGA 0/1 at Week 16 – key secondary PASI 75 at Week 16 – key secondary	FAS	NRI+CMH	
PASI 90 at Week 16 – key secondary	FAS	NRI+CMH	
Change from baseline in PSSD symptom score at Week 16 – key secondary	FAS	mBOCF+ANCOVA	
ss-PGA 0/1 among subjects with a baseline ss-PGA ≥ 3 at Week 16 – key secondary	FAS	NRI+CMH	
sPGA 0 at Week 16 – key secondary	FAS	NRI+CMH	
PSSD symptom score of 0 at Week 16 – key secondary	FAS	NRI+CMH	
sPGA 0/1 at Week 24 – key secondary PASI 75 at Week 24 – key secondary	FAS	NRI+CMH	
PASI 90 at Week 24 – key secondary	FAS	NRI+CMH	
sPGA 0/1 at Week 52 and Week 24 – key secondary	FAS	NRI+CMH	
PASI 75 at Week 52 and Week 24 – key secondary	FAS	NRI+CMH	

List of Planned Analyses: Comparisons of BMS-986165 6 mg QD vs. Apremilast			
Measure of Interest	**Population**	**Primary Analysis at Week 16/24/52**	[redacted]
PASI 90 at Week 52 and Week 24 – key secondary	FAS	NRI+CMH	

Appendix 2 Summary of Efficacy Assessments

Assessment	Outcome Measure	Endpoint	Comparison	Analyses
sPGA (Sec 8.2.1.1)	sPGA 0/1 with at least 2 point improvement from baseline	W16	BMS vs. PBO	CMH, sensitivity analyses including tipping point analysis, multiple imputation (Sec 6.1.1.1-3)
			BMS vs. apremilast	CMH (Sec 6.1.2.1)
		W24/52	BMS vs. apremilast	CMH (Sec 6.1.2.1)
		[redacted]		
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
		Baseline->W52 over time	BMS vs. apremilast	Binary GEE & CMH by Week (Sec 6.1.5.1)
	sPGA 0	W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
			BMS vs. apremilast	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
		[redacted]		
	[redacted]			
PASI (Sec 8.2.1.2)	PASI 75	W16	BMS vs. PBO	CMH, sensitivity analyses including tipping point analysis, multiple imputation (Sec 6.1.1.1-3)
			BMS vs. apremilast	CMH (Sec 6.1.2.1)
		W24/52	BMS vs. apremilast	CMH (Sec 6.1.2.1)
		[redacted]		

Assessment	Outcome Measure	Endpoint	Comparison	Analyses
		[redacted]		
		Baseline->Week 16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
		Baseline->Week 52 over time	BMS vs. apremilast	Binary GEE & CMH by Week (Sec 6.1.5.1)
	PASI 90	W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
			BMS vs. apremilast	CMH (Sec 6.1.2.1)
		W52	BMS vs. apremilast	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
		Baseline->W52 over time	BMS vs. apremilast	Binary GEE & CMH by Week (Sec 6.1.5.1)
	PASI 100	W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
			BMS vs. apremilast	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
		Baseline->W52 over time	BMS vs. apremilast	Binary GEE & CMH by Week (Sec 6.1.5.1)
	[redacted]			
[redacted]				

Assessment	Outcome Measure	Endpoint	Comparison	Analyses
[redacted]				
ss-PGA (Sec 8.2.1.4)	ss-PGA 0/1 among subjects with a baseline ss-PGA ≥3	W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
			BMS vs. apremilast	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
[redacted]				
PGA-F (Sec 8.2.1.6)	PGA-F 0/1 among subjects with a baseline PGA-F score ≥3	W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
[redacted]				
pp-PGA (Sec 8.2.1.8)		W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	CMH by Week (Sec 6.1.5.1)

Assessment	Outcome Measure	Endpoint	Comparison	Analyses
	pp-PGA 0/1 among subjects with a baseline pp-PGA ≥3	[redacted]		
PSSD / Symptom Score (Sec 8.2.2.1)	PSSD/Symptom 0 among subjects with baseline PSSD/symptom score ≥1	W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
			BMS vs. apremilast	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)

Assessment	Outcome Measure	Endpoint	Comparison	Analyses
[redacted]				
DLQI (Sec 8.2.2.4)	DLQI 0/1 among subjects with a baseline DLQI score ≥2	W16	BMS vs. PBO	CMH (Sec 6.1.2.1)
		Baseline->W16 over time	BMS vs. PBO	Binary GEE & CMH by Week (Sec 6.1.5.1)
[redacted]				

Assessment	Outcome Measure	Endpoint	Comparison	Analyses
[redacted]				
PASE (Sec 8.2.2.11)	PASE	Screening only		No Analysis