Statistical Analysis Plan

Protocol Title:	A PHASE 3, MULTI-CENTER, RANDOMIZED, DOUBLE- BLIND, PLACEBO-CONTROLLED STUDY TO ASSESS THE EFFICACY AND SAFETY OF APREMILAST (CC-10004) IN PEDIATRIC SUBJECTS FROM 6 THROUGH 17 YEARS OF AGE WITH MODERATE TO SEVERE PLAQUE PSORIASIS
Short Protocol Title:	Not Applicable
Protocol Number:	20200056 (CC-10004-PPSO-003)
NCT Number:	NCT03701763
Authors:	[redacted], [redacted], [redacted]
Sponsor:	Amgen Inc. One Amgen Center Drive, Thousand Oaks, CA 91320, United States
SAP Date:	**Document Version** / **Date** **Original (v2.0)** / 20 December 2021

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Celgene Version 0.2	21Aug2018	Original
Amgen Version 1.0	31Aug2021	Sample size reduction to at least 180 subjects to reach committed timelines with regulatory authorities while keeping sufficient power.
Amgen Version 2.0	20Dec2021	Sample size was reverted from at least 180 randomized subjects to the original protocol goal of at least 230 randomized subjects based on enrollment progress for the study and advice from the United States Food and Drug Administration. The other updates are presented in the following page of Summary of Changes.

Summary of Changes

1. **Updated the document to align with the Amgen template for statistical analysis plans and adjusted the section numbers accordingly.**
2. **Reworded the Hypotheses as follows:**

   "The primary hypothesis for this study is: Apremilast is superior to placebo in subjects with moderate to severe psoriasis as measured by proportion of subjects with a sPGA score of clear (0) or almost clear (1) with at least a 2-point reduction from baseline at Week 16.

   The major secondary hypothesis for this study is: Apremilast is superior to placebo in subjects with moderate to severe psoriasis as measured by proportion of subjects achieving at least a 75% reduction in PASI (PASI-75) from baseline at Week 16."

   Versus the original text:

   "The null hypothesis is that the effects of the two treatment arms (i.e., placebo vs. apremilast) have no difference" in Page 27.

   Affected section: Section 2.2
3. **Added "Data Screening and Acceptance" details following the Amgen smart template.**

   Affected section: Section 8.1, 8.2, 8.4, 8.6, 8.7
4. **Added "Withdrawal by parent/guardian" and "Discontinued reason due to COVID-19 related control measures" in the categories of disposition.**

   Affected section: Section 9.4
5. **Updated the MedDRA version and WHO DD version from 20.0 and March 2018 to 24.0 or higher and March 2020 or later, respectively.**

   Affected section: Section 9.5.3, 9.5.4, 9.5.5, 9.8.1, 9.8.7, 9.8.8
6. **Added Analyses of COVID-19 impacted data: "Analyses will be performed to assess the impact of COVID-19 on the amount of data collected for the study."**

   Affected section: Section 9.9.1

Table of Contents

List of Table

List of Figures

List of Appendices

List of Abbreviations

Abbreviation	Explanation
AE	Adverse Event
ANCOVA	Analysis of Covariance (model)
ATC	Anatomical-Therapeutic-Chemical
BID	Twice Daily
BMI	Body Mass Index
BSA	Body Surface Area (%) Affected by Psoriasis
CDC	Centers for Disease Control and Prevention
CDLQI	Children's Dermatological Life Quality Index
CI	Confidence Interval
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus Disease 2019
CRF or eCRF	(electronic) Case Report Form
CRO	Contract Research Organization
CRP	Clinical Research Physician
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide Severity Rating Scale
DMC	Data Monitoring Committee
FCBP	Female of Childbearing Potential
FDLQI	Family Dermatological Life Quality Index
HRQoL	Health-Related Quality of Life
IP	Investigational Product
IRT	Interactive Response Technology
ITT	Intent-to-Treat
LOCF	Last Observation Carried Forward
LS Mean	Least Square Mean
MCID	Minimal Clinically Important Difference
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
NRI	Non-responder Imputation

NRS	Numeric Rating Scale
PASI	Psoriasis Area Severity Index
PBO	Placebo
PD	Pharmacodynamic
PG	Pharmacogenetic
PK	Pharmacokinetic
PP	Per-Protocol (population)
PRO	Patient Reported Outcome
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
ScPGA	Scalp Physician Global Assessment
SD or STDEV	Standard Deviation
SE	Standard Error
SI	Standard International (unit)
sPGA	Static Physician Global Assessment
sPGA-G	Static Physician Global Assessment of Genitalia
SOC	System Organ Class
TEAE	Treatment-emergent Adverse Event
USA	United States of America
WHODD	World Health Organization Drug Dictionary

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol for apremilast (CC-10004, AMG 407) Study CC-10004-PPSO-003 (20200056), dated 17 Dec. 2021. The scope of this plan includes the primary analysis and the final analysis that are planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified.

2. Objectives, Endpoints and Hypotheses

2.1 Objectives and Endpoints/Estimands

Objectives	Endpoints
Primary	
• To evaluate the clinical efficacy of apremilast compared with placebo in children and adolescents (ages 6 through 17 years) with moderate to severe plaque psoriasis	• sPGA 0/1: Proportion of subjects with an sPGA score of clear (0) or almost clear (1) with at least a 2-point reduction from baseline at Week 16
Secondary	
• To evaluate the safety and tolerability of apremilast compared with placebo, in children and adolescents (ages 6 through 17 years) with moderate to severe plaque psoriasis • To evaluate the effect of apremilast compared with placebo on health-related quality of life (HRQoL)	• PASI-75: Proportion of subjects who achieve at least a 75% reduction in PASI (PASI-75) from baseline at Week 16 • PASI-50: Proportion of subjects who achieve at least a 50% reduction in PASI (PASI-50) from baseline at Week 16 • PASI: Percent change from baseline in total PASI score at Week 16 • BSA: Percent change from baseline in affected BSA at Week 16 • CDLQI (0/1): Proportion of subjects who achieve CDLQI (0/1) at Week 16 • CDLQI: Change from baseline in CDLQI score at Week 16 • Adverse events: Type, frequency, severity, and relationship of AEs to IP for study duration (up through 14-week follow-up) • Diarrhea: Stool Diary (frequency, duration and associated symptoms)

	for Study duration (up through 4-week follow-up) • Depression, suicidal thoughts and behavior: Columbia-Suicide Severity Rating Scale (C-SSRS) questionnaire completed at screening and each visit thereafter for study duration (up through Week 52) • Growth and Development: - Tanner Staging of sexual development performed at beginning and end of study (Week 52 or ET) - Body weight, height and BMI for study duration • Psoriasis Flare for study duration (while taking study medication) • Psoriasis Rebound in 14-week follow-up period

Exploratory	
• To evaluate the effect of apremilast on plaque psoriasis • To evaluate the effect of apremilast on plaque psoriasis of the scalp • To evaluate the effect of apremilast on genital psoriasis • To evaluate the effect of apremilast on itch over the whole body caused by plaque psoriasis • To evaluate the effect of apremilast on HRQoL of family members [redacted]	• PASI-50: Proportion of subjects who achieve PASI-50 at all visits • PASI-75: Proportion of subjects who achieve PASI-75 at all visits • PASI-90: Proportion of subjects who achieve PASI-90 at all visits • PASI: Percent change from baseline in total PASI score at all visits • sPGA (0/1): Proportion of subjects with sPGA of clear (0) or almost clear (1) with at least a 2-point reduction from baseline at all visits • BSA: Percent change from baseline in affected BSA at all visits • ScPGA: Proportion of subjects with scalp psoriasis with improvement of ScPGA of clear (0) or almost clear (1) at select visits • Modified sPGA-G: Proportion of subjects with genital psoriasis with

[redacted]	improvement of sPGA-G of clear (0) or almost clear (1) at select visits • Modified Whole Body Itch NRS: Proportion of subjects with ≥ 4-point reduction (improvement) from Baseline in the whole body itch NRS score at select visits • Modified Whole Body Itch NRS: Change from Baseline in the whole body itch NRS score at select visits • CDLQI (0/1): Proportion of subjects who achieve CDLQI (0/1) at select visits • CDLQI: Change from baseline in CDLQI score at select visits • FDLQI (0/1): Proportion of subjects achieving FDLQI (0/1) at select visits • FDLQI: Change from baseline in FDLQI score at select visits [redacted]

2.2 Hypotheses and/or Estimations

The primary hypothesis for this study is: Apremilast is superior to placebo in subjects with moderate to severe psoriasis as measured by proportion of subjects with a sPGA score of clear (0) or almost clear (1) with at least a 2-point reduction from baseline at Week 16.

The major secondary hypothesis for this study is: Apremilast is superior to placebo in subjects with moderate to severe psoriasis as measured by proportion of subjects achieving at least a 75% reduction in PASI (PASI-75) from baseline at Week 16.

3. Study Overview

3.1 Study Design

This is a Phase 3, multicenter, randomized, placebo-controlled, double-blind study of the efficacy and safety of apremilast (CC-10004) in pediatric subjects with moderate to severe plaque psoriasis.

At least 230 pediatric subjects (ages 6 through 17 years) will be randomized 2:1 to receive either apremilast or placebo for the first 16 weeks. Randomization to apremilast arm or placebo arm will be stratified by age group (6-11 yrs or 12-17 yrs). A minimum of 75 subjects will be randomized in each age group. Treatment will be assigned by baseline weight with subjects 20 kg to < 50 kg receiving apremilast 20 mg BID or placebo BID and subjects ≥ 50 kg receiving apremilast 30 mg BID or placebo BID. After all subjects have completed Week 16 (Visit 7), or discontinued from the study, a Week 16 database lock will be performed; the primary data analysis will be conducted and a Week 16 clinical study report will be generated. Study investigators and subjects will remain blinded to initial treatment assignments until the final database lock at the conclusion of the study. The blind also should be maintained for persons responsible for the ongoing conduct of the study. Blinded persons may include but are not limited to: clinical research physician, clinical research scientist, clinical trial manager, study statistician, data manager, statistical programmer, and clinical research associate. At the end of the study, after all subjects have either completed Week 52 (Visit 16) and entered the Long-term Study, or completed the Observational Follow-up Phase, a final analysis will be performed and a final clinical study report will be generated.

Subjects must have a diagnosis of chronic, stable plaque psoriasis at least 6 months prior to screening and which is considered inadequately controlled by or inappropriate for topical therapy to qualify for this study. In addition, the subjects must have a Psoriasis Area Severity Index (PASI) score ≥ 12, body surface area (BSA) involvement ≥ 10% and static Physician Global Assessment (sPGA) ≥ 3 (moderate to severe) at screening and baseline (Week 0).

The study will consist of four phases:

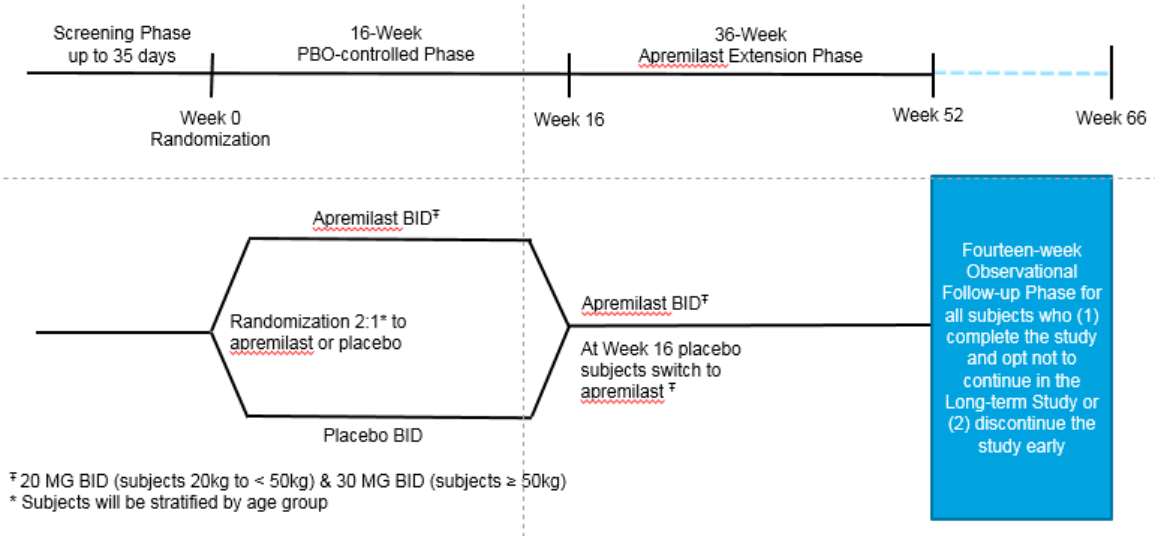
- Screening Phase – up to 35 days
- Double-blind, Placebo-controlled Phase – Weeks 0-16
    - Subjects will be randomly assigned in a 2:1 ratio to weight-based apremilast or placebo.

- From Week 8 through Week 16, any subject with a PASI increase ≥ 50% from baseline will be eligible to commence treatment with moderate-to-high potency topical steroid preparations (early escape) and continue with randomized IP.

- Apremilast Extension Phase – Weeks 16-52
  - Placebo subjects will be switched at Week 16 to receive apremilast 20 mg BID or 30 mg BID, according to baseline weight. All other subjects will continue to receive either apremilast 20 mg BID or apremilast 30 mg BID based on their original dosing assignment.

- Observational Follow-up Phase – 14 weeks
  - All eligible subjects who complete the Apremilast Extension Phase may opt to enroll in a separate Long-term Study (for up to 4 years or until approval, whichever comes first). Subjects who choose not to participate in the Long-term Study, or discontinue the study early, should return for observational follow-up visits four, eight, and fourteen weeks after the last dose of study drug.

Figure 1: Study Design



3.2 Sample Size

At least 230 pediatric subjects (ages from 6 through 17 years) will be randomized to receive either apremilast or placebo (2:1) at Week 0. The sample size estimation is based on the results of the adult Phase 3 and 3b studies with apremilast (CC-10004-PSOR-008, PSOR-009 and PSOR-010) which demonstrated positive treatment effects

between apremilast and placebo in the proportion of subjects achieving sPGA response and PASI-75 response at Week 16.

With a total of 230 patients and a randomization ratio of 2:1, the study will have 90% power to detect a 15% difference (20% versus 5%) between the apremilast arm and the placebo arm for the proportion of subjects achieving sPGA response at Week 16, based on a Chi-square test at a two-sided significance level of 0.05. The study will also have more than 95% power to detect a 20% difference (30% versus 10%) between apremilast and placebo for the proportion of subjects achieving PASI-75 response at Week 16.

3.3 Adaptive Design

Not applicable.

4. Covariates and Subgroups

4.1 Planned Covariates

The study will be stratified by age group (6 to 11 years or 12 to 17 years). Of the analyses, the CMH test for the binary endpoints and the ANCOVA model for the continuous endpoints will be adjusted for the baseline value.

4.2 Subgroups

Subgroup analysis will be carried out for sPGA and PASI-75 responses at Week 16 in the Placebo-controlled Phase based upon baseline demographics and disease characteristics. These two endpoints will also be evaluated in patients using or not using low potency topical corticosteroids in the phase as two subgroups. Summary and analysis will be based on ITT population and missing values will be imputed using MI method. Treatment difference of response rates will be reported with two-sided 95% CI.

The following subgroup variables will be used:

- Sex (Male, Female)
- Race (White, Non-white)
- Age category (6-11 years, 12-17 years)
- Baseline weight category (≥ 20 kg to < 50 kg, ≥ 50 kg)
- Age and baseline weight category
    - 6-11 years and ≥ 20 kg to < 50 kg
    - 6-11 years and ≥ 50 kg
    - 12-17 years and ≥ 20 kg to < 50 kg

  - 12-17 years and ≥ 50 kg

- Baseline BMI category (Underweight, Healthy Weight, Overweight, Obesity)
- Geographical region (USA, Canada, Europe, Rest of the World)
- Duration of plaque psoriasis categories
- Baseline sPGA score 3 (moderate), 4 (severe)
- Baseline PASI score categories
- Baseline BSA (%) categories
- Baseline CDLQI total score categories
- Baseline whole body itch NRS categories
- Baseline ScPGA categories
- Baseline sPGA-G categories
- Prior phototherapies (Used or Never used)
- Prior conventional systemic therapies (Used or Never used)
- Prior biologic therapies (Used or Never used)
- Prior systemic therapies (Used or Never used)

5. Definitions

5.1 Baseline Definitions

For efficacy analysis and summary of baseline disease characteristics data, baseline is defined as the last value measured prior to or at the randomization date.

For safety analyses in the Placebo-controlled Phase, baseline will be relative to the first dose date following randomization at Week 0. It is the last value measured prior to or on the day of the first dose of IP.

For safety analyses in Apremilast Exposure Period, baseline will be relative to the first apremilast dose date at Week 0 for subjects initially randomized to apremilast or Week 16 for subjects initially randomized to placebo and switched to apremilast in the Extension Phase. It is the last value measured prior to or at the first apremilast dose date.

5.2 Time Points

Time points in all analyses are based on the remapped visits/study weeks using the following visit mapping algorithm, which may or may not be the same as the visits/study weeks as recorded in the database.

Data collected in the placebo-controlled phase and apremilast extension phase will be summarized based on the weeks specified in protocol; the observational Follow-up Phase will be summarized separately using the three follow-up visits, i.e., 4-Week Follow-up, 8-Week Follow-up, 14-Week Follow-up.

The details are specified in Appendix B.

5.3 Derivation of Efficacy Endpoints

The derivation of efficacy endpoints is described below in separate sections. Baseline definition for all efficacy endpoints is given in Section 5.1. Change from baseline is calculated as post-baseline visit value minus the baseline value. Percent change from baseline is defined as 100* Change from baseline/Baseline value (%). Handling of time points is described in Section 5.2.

5.3.1 Static Physician Global Assessment (sPGA)

The sPGA is the assessment of whole body psoriasis by the investigator of the overall disease severity at the time of evaluation. The sPGA is a 5-point scale ranging from clear (0), almost clear (1), mild (2), moderate (3), to severe (4), incorporating an assessment of the severity of the three primary signs of the disease: erythema, scaling and plaque elevation.

The sPGA response at a post-baseline visit is defined as achieving sPGA score of clear (0) or almost clear (1) with at least a 2-point reduction from baseline.

5.3.2 Psoriasis Area and Severity Index (PASI)

The PASI is a measure of psoriatic disease severity taking into account qualitative lesion characteristics (erythema, thickness, and scaling) and degree of skin surface area involvement on defined anatomical regions. PASI scores range from 0 to 72, with higher scores reflecting greater disease severity. Erythema, thickness, and scaling are scored on a scale of 0 (none) to 4 (very severe) on 4 anatomic regions of the body: head, trunk, upper limbs, and lower limbs. Degree of involvement on each of the 4 anatomic regions is scored on a scale of 0 (no involvement) to 6 (90% to 100% involvement). The total qualitative score (sum of erythema, thickness, and scaling scores) is multiplied by the degree of involvement for each anatomic region and then multiplied by a constant.

These values for each anatomic region are summed to yield the PASI score. The PASI score will be set to missing if any severity score or degree of involvement is missing.

PASI scores will be presented with one digit after the decimal point. For post-baseline visits, PASI score change and percent change from baseline will be derived. PASI score percent change will be rounded to integers.

For post-baseline visit, a subject is classified as having achieved PASI-75 (PASI-75 response) if the PASI score is reduced by at least 75% from baseline, which is equivalent to a percent change from baseline ranging from −100% to −75%. PASI-50 and PASI-90 are similarly defined as PASI score reduction from baseline by at least 50% or 90%, respectively.

5.3.3 Body Surface Area (BSA)

BSA is a measurement of involved skin over the whole body. The overall BSA affected by psoriasis is estimated based on the palm area of the subject's hand. The surface area of the whole body is made up of approximately 100 palms or "handprints" (each entire palmar surface or "handprint" equates to approximately 1% of total body surface area).

BSA scores will be presented with one digit after the decimal point. For post-baseline visits, BSA score change and percent change from baseline will be derived. BSA score percent change will be rounded to integers.

5.3.4 Scalp Physician Global Assessment (ScPGA)

The ScPGA is the Investigator's assessment of severity of psoriasis of the scalp, if present at Baseline. The ScPGA is a 5-point scale ranging from clear (0), almost clear (1), mild (2), moderate (3), to severe (4).

The ScPGA response at a post-baseline visit is defined as achieving ScPGA score of clear (0) or almost clear (1).

5.3.5 Modified Static Physician Global Assessment of Genitalia (sPGA-G)

The modified sPGA-G is the Investigator's assessment of severity of psoriasis of the genitals, if present at Baseline. It is a 5-point scale ranging from clear (0), almost clear (1), mild (2), moderate (3), to severe (4).

The modified sPGA-G response at a post-baseline visit is defined as achieving sPGA-G score of clear (0) or almost clear (1).

5.3.6 Modified Whole Body Itch Numeric Rating Scale (NRS) Assessment

The modified whole body itch NRS assessment is a tool designed to measure the amount of whole body itch due to psoriasis by circling a number on a scale from 0 to 10. Subjects will be asked to assess their worst level whole body itch in the past 24 hours and select a number on a scale of 0 to 10, where “0” represents no itching, and “10” represents the worst itch imaginable.

For a post baseline visit, the whole body itch NRS response is defined as ≥ 4 points reduction (improvement) from baseline in the NRS score. Change from baseline in the NRS score will also be derived.

5.3.7 Children’s Dermatology Life Quality Index (CDLQI)

The Children’s Dermatology Life Quality Index questionnaire is designed for use in children, i.e. patients from age 4 to age 16. It is self-explanatory and can be simply handed to the patient who is asked to fill it in with the help of the child’s parent or guardian. The instrument contains 10 questions, which have scores ranging from Very much (3), Prevented school (3) for Question 7, Quite a lot (2), Only a little (1), to Not at all (0).

The CDLQI is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score, the more quality of life is impaired. If one of the 10 items is left unanswered, it is scored 0 and the total scores are summed and expressed as usual out of a maximum of 30. If two or more of the ten items are left unanswered, the total score will be left missing. If both parts of question 7 are completed the higher of the two scores should be counted. The severity banding for CDLQI total scores are:

- 0-1 = no effect on child’s life
- 2-6 = small effect
- 7-12 = moderate effect
- 13-18 = very large effect
- 19-30 = extremely large effect

For CDLQI total score, change form baseline will be derived, where change = visit value − baseline value. For a post-baseline visit, CDLQI (0/1) response is defined as achieving CDLQI total scores 0 or 1.

5.3.8 Family Dermatology Life Quality Index (FDLQI)

The Family Dermatology Life Quality Index is a questionnaire designed for adult (more than 16 years of age) family members or partners of patients with any skin disease. The FDLQI total score is calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score, the more quality of life is impaired.

If one question is left unanswered then it is scored 0 and the scores are summed and expressed as usual out of a maximum of 30. If two or more questions are left unanswered then the questionnaire is not scored.

For FDLQI total score, change from baseline will be derived, where change = visit value − baseline value. For a post-baseline visit, FDLQI (0/1) response is defined as achieving FDLQI total scores 0 or 1.

5.4 Derivation of Safety Endpoints

Baseline definition for all safety endpoints is given in Section 5.1. Change from baseline is calculated as post-baseline visit value minus the baseline value. Handling of time points is described in Section 5.2.

5.4.1 Treatment-emergent Adverse Event

An AE is a treatment-emergent AE (TEAE) if the AE start date is

- On or after the date of the first dose of IP and no later than 28 days after the last dose of IP for subjects who have completed study treatment or have discontinued early by the time of database cut, or
- On or after the date of the first dose of IP for subjects who are ongoing with study treatment at the time of database cut.

If the treatment-emergent status of an AE is unclear due to a missing/incomplete start date, it will always be considered treatment-emergent, unless shown otherwise by data. Date imputation rules for missing AE start dates are described in Appendix C.

For the Placebo-controlled Phase (Weeks 0-16), TEAEs include adverse events started on or after the first IP dose date and prior to the first dose date of the Apremilast Extension Phase for subjects who also received treatment in the Apremilast Extension Phase or within 28 days of the last dose date for subjects who received treatment only in the Placebo-controlled Phase.

For the Apremilast Extension Phase (Weeks 16-52), TEAEs include adverse events started on or after the first dose date of the Apremilast Extension Phase and within 28 days of the last dose date.

For the Apremilast-exposure Period, TEAEs include adverse events started on or after the first Apremilast dose date and within 28 days of the last Apremilast dose date.

Adverse event started more than 28 days after the last dose of IP will not be considered as a treatment-emergent AE.

5.4.2 Treatment-emergent Adverse Events Leading to Drug Withdrawal, Leading to Drug Interruption, and Leading to Death, and Drug-related Treatment-emergent Adverse Events

A TEAE leading to drug withdrawal is a TEAE for which the investigator indicates that the action taken with respect to IP is withdrawn permanently. A TEAE leading to drug interruption is a TEAE for which the investigator indicates that the action taken with respect to IP is interrupted. A TEAE leading to death is a TEAE for which the outcome is fatal. Relationship to IP is based on the investigator's causality judgment; that is, a drug-related AE is an AE indicated by the investigator to have a suspected relationship to IP.

5.4.3 Stool Diary for Diarrhea

Subjects and their parent/guardian will be supplied with paper diaries that will be filled out daily to record and describe any diarrhea, including duration, frequency, treatment, and associated symptoms. Subjects and their parent/guardian will complete the Stool Diary every day beginning after the first dose of apremilast, up to the 4-week post last dose follow-up visit.

5.4.4 Psychiatric Evaluation

Subjects will complete the Columbia-Suicide Severity Rating Scale (C-SSRS) assessment. This questionnaire is suitable for assessment of suicidal ideation and behavior in clinical and research settings. The assessment will be completed at screening, baseline, and at post-baseline visits.

Any subject that answers ‘Yes’ to any question on the C-SSRS at any time during the study will be immediately withdrawn from study treatment.

5.4.5 Psoriasis Flare and Psoriasis Rebound

Psoriasis flare was considered an AE (and was to be recorded as an AE) and represents an atypical or unusual worsening of disease during treatment. It was defined as a sudden intensification of psoriasis requiring medical intervention or a diagnosis of new

generalized erythrodermic, inflammatory, or pustular psoriasis. A more typical, gradual worsening of plaque psoriasis was not to be recorded as an AE.

Rebound was considered as AE and was defined as a severe and sudden worsening of disease that occurred after treatment had been discontinued. This exacerbation is characterized by PASI ≥ 125% of baseline or a new generalized pustular, erythrodermic, or more inflammatory psoriasis after stopping therapy.

Psoriasis adverse events will be selected as those containing the word “psoriasis” in the coded preferred term. A psoriasis flare is defined as a psoriasis adverse event started during treatment with IP. A psoriasis rebound is defined as a psoriasis adverse event started after the last dose of IP.

5.4.6 Body Weight, Height and Body Mass Index (BMI)

Body weight (kg) and Height (cm) will be measured and recorded at study visits. The BMI (kg/m^2) is calculated as BMI = Weight / $(Height/100)^2$. The variables include the observed value, change and percent change from baseline over time.

It is now common practice to express child growth status in the form of the BMI standard deviation scores (Z-score). The LMS method provides a way of obtaining normalized growth centile standards. It assumes that the data can be normalized by using a power transformation, which stretches one tail of the distribution and shrinks the other, removing the skewness. The LMS parameters are the power in the Box-Cox transformation (L), the median (M), and the generalized coefficient of variation (S).

For a given value of BMI of a boy or girl with known age, the corresponding z-score (Z) can be calculated as:

$Z = [(BMI/M)^L – 1]/LS,$

where the coefficients L, M, S are from the BMI percentile values in Data Tables of BMI-for-age Charts (CDC 2000). The tables list L, M, S values for boys or girls with age ranging from 2 to 20 years old.

BMI are classified using z-score (Z) as:

- Underweight: Percentile < 5
- Healthy Weight: Percentile ≥ 5 to < 85
- Overweight: Percentile ≥ 85 to < 95
- Obesity: Percentile ≥ 95

5.4.7 Vital Signs

Vital signs include:

- Observed value and change from baseline over time in vital signs (temperature, pulse, and blood pressure)
- Shifts from baseline to post-baseline timepoints and to the worst post-baseline value in terms of normal/abnormal in pulse and blood pressure (normal ranges are defined as: 60-100 beats/minute for pulse, 90-140 mmHg for systolic blood pressure, and 60-90 mmHg for diastolic blood pressure)

5.4.8 Clinical Laboratory Evaluations

Laboratory evaluations include:

- Observed value and change from baseline over time in hematology and serum chemistry parameters
- Laboratory marked abnormalities
- Shifts from baseline to post-baseline time points and to the worst post-baseline value in terms of normal/abnormal (low or high) in hematology and serum chemistry parameters

6. Analysis Sets

6.1 Intent-to-treat Population

The intent-to-treat (ITT) population will consist of all subjects who are randomized regardless of whether or not the subject received IP. Subjects will be included in the treatment group to which they are randomized.

6.2 Safety Analysis Set

6.2.1 Placebo-controlled Phases (Weeks 0 to16)

The safety population will consist of all subjects who are randomized and received at least one dose of investigational product (IP). Subjects will be included in the treatment group corresponding to the IP they actually received (apremilast or placebo) for the analyses and summaries using the safety population.

6.2.2 Apremilast Extension Phase (Weeks 16 to 52)

Safety analyses will be based on subjects who are treated in the phase, ie, who receive at least one dose of apremilast in the extension phase. At Week 16, subjects will continue apremilast treatment or will switch to apremilast treatment from placebo.

6.2.3 Apremilast-exposure Period

The safety analyses for the Apremilast-exposure Period will be based on the apremilast subjects as treated population, which includes all subjects who are randomized to (at Week 0/Visit 2) or treated with (at Week 16/Visit 7) apremilast, and receive at least one dose of apremilast after randomization or Week 16.

6.3 Per Protocol Set(s)

The per protocol (PP) population will consist of all subjects included in the ITT population who receive at least one dose of IP, have both baseline and at least one post-treatment sPGA assessment, and have no important protocol deviations (IPD) which may affect efficacy assessments in the Placebo-controlled Phase.

6.4 Health-related Quality-of-Life or Health Economics Analyses Set(s)

Not applicable.

[redacted]

6.6 Interim Analyses Set(s)

Not applicable.

6.7 Study-specific Analysis Sets

Not applicable.

7. Planned Analyses

7.1 Interim Analysis and Early Stopping Guidelines

No interim analysis will be conducted.

After all subjects have completed Week 16 (Visit 7), or discontinued from the study, a Week 16 database lock will be performed; the primary data analysis will be conducted and a Week 16 Clinical Study Report will be generated. Study investigators and subjects will remain blinded to treatment assignments until the final database lock at the conclusion of the study. The blind also should be maintained for persons responsible for the ongoing conduct of the study. Blinded persons may include but are not limited to: clinical research physician, clinical research scientist, clinical trial manager, study statistician, data manager, programmer, and clinical research associate. At the end of the study, after all subjects have either completed the Observational Follow-up Phase, or have been discontinued from the Apremilast Extension Phase (Weeks 16 to 52) or the

Observational Follow-up Phase, the final analysis will be performed and a final Clinical Study Report will be generated.

7.2 Primary Analysis

The primary analysis will be conducted once all subjects have completed week 16 or terminated early from the study. At this time, the data will be cleaned, processed, and extracted, and restricted. The endpoints will be assessed at the time of the primary analysis.

To maintain the blind at the site and subject level, the individual subject treatment assignments will not be revealed to the Investigators until after the final database lock following the study completion.

7.3 Final Analysis

The final analysis will be conducted after study completion (end of study for the last subject) is reached and all study data are collected.

7.4 Analysis Phases or Periods

For efficacy analysis, different phases will be used. For safety analysis, different phases and Apremilast-exposure Period will be used.

7.4.1 Analysis Phases

Per protocol specification, data summary and analysis will be provided for the following phases.

- Placebo-controlled Phase – Weeks 0 to16

This phase starts on the day of randomization (Week 0/Visit 2), and stops on either: (1) the day the first IP for the next phase is dispensed at Week 16/Visit 7; or (2) the day of the discontinuation visit if the subject discontinued prior to or at Week 16/Visit 7; or (3) the last known study day if the subject is lost to follow-up prior to Week 16/Visit 7 during the phase.

For safety analysis in subjects who continued treatment in Apremilast Extension Phase (Weeks 16 to 52), the phase stopped one day prior to the first dose date in Apremilast Extension Phase.

- Apremilast Extension Phase – Weeks 16 to 52

This phase starts on the next day the first IP is dispensed for the phase at Week 16/Visit 7, and stops on either: (1) the day of Week 52/Visit 16; or (2) the day of the discontinuation visit if the subject discontinued prior to or at Week 52/Visit 16; or (3) the

last known study day if the subject lost to follow-up prior to Week 52/Visit 16 during the phase.

For safety analysis, the phase started at the first dose date in the phase.

- Observational Follow-up Phase – 14 weeks

Subjects who choose not to participate in a separate Long-term Study (for up to 4 years or until approval, whichever comes first), or discontinue the study early, should return for observational follow-up visits four, eight and fourteen weeks after the last dose of study drug. The 14 weeks Observational Follow-up Phase starts at the next day after the completion or discontinuation visit and stops at the last follow-up visit or the last assessment date.

7.4.2 Apremilast-exposure Period for Safety Analysis

In addition to the above defined phases, Apremilast-exposure Period will be used for safety analyses.

- Apremilast-exposure Period:

This period starts on the date of either: (1) the first dose of IP following randomization (Week 0/Visit 2) for subjects who are treated with apremilast from Week 0; or (2) the first dose of IP from the IP dispensed at Week16/Visit 7 for subjects who were originally treated with placebo and are treated with apremilast at Week 16.

This period stops on either: (1) data cut-off date; or (2) the day of the treatment discontinuation if the subject discontinued prior to or at Week 52/Visit 16; or (3) the last known study day if the subject lost to follow-up prior to Week 52/Visit 16; or (4) Week 52/Visit 16 visit date.

8. Data Screening and Acceptance

8.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

8.2 Data Handling and Electronic Transfer of Data

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. This study will use the RAVE database.

8.3 Handling of Missing and Incomplete Data

Missing values at Week 16 for efficacy endpoint will be imputed using the multiple imputation (MI) method based on similar subjects who remained in the study as the primary method. Sensitivity analysis will be conducted to account for missing data using the last observation carried forward (LOCF) method, the non-responder imputation (NRI) method and the tipping point analyses. A sensitivity analysis using the PP population will also be performed.

The handling of incomplete and partial dates for adverse events and concomitant medications etc. will be described in Appendix C.

8.4 Detection of Bias

The measures to Minimize Bias in this study are Randomization and Blinding. It is not expected that any study conduct procedures or statistical analysis will incur bias in the study results or conclusions. The potential sources of bias in this study are:

a) Inadvertent breaking of the blind

Break of blind should be rare and anticipated only if it is necessary for safety and treatment reasons. Should this happen in more than 5% of the subjects in the study, a sensitivity analysis may be conducted for the primary endpoint by excluding those patients who were unblinded.

b) Protocol deviations

A subject list of important protocol deviations will be finalized prior to the database lock and unblinding. The impact of protocol deviations on the study results will be investigated and a sensitivity analysis may be conducted if deemed necessary by comparing the results of the primary endpoint analysis on subjects with and without protocol deviations.

8.5 Outliers

Various methods, including univariate summaries, histograms, scatter plots, box plots, and line graphs, may be used to identify outliers in key safety and efficacy variables.

The primary analyses will include outliers in the data. Sensitivity analyses may be undertaken if extreme outliers for a variable are observed.

8.6 Distributional Characteristics

Where appropriate, the assumptions underlying the proposed statistical methodologies will be assessed. If required, data transformations or alternative non-parametric methods of analyses will be utilized.

8.7 Validation of Statistical Analyses

Programs will be developed and maintained, and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software; for example, the SAS System version 9.4 or later.

9. Statistical Methods of Analysis

9.1 General Considerations

The intent-to-treat principle will be used in statistical analyses for efficacy endpoints. For the Placebo-controlled Phase (Weeks 0-16), efficacy evaluations will be conducted using the intent-to-treat (ITT) population defined as all randomized subjects. Statistical comparisons will be made between the two treatment arms (placebo or apremilast). The apremilast arm will include both 20 mg BID and 30 mg BID, since subjects will receive weight-based apremilast doses which will provide comparable exposure to the adult subjects treated with 30 mg BID. For the Apremilast Extension Phase (Weeks 16-52), efficacy evaluations will be conducted using subjects who entered the phase from both treatment arms.

Descriptive statistics (n, mean, SD, median, Q1, Q3, min, max) will be presented for appropriate endpoints at specified time points. Specifically, for continuous variables, descriptive statistics for assessment values and change (plus percentage change when specified) from baseline will be provided. Binary variables will be summarized with frequency tabulations.

Statistical comparisons will be made between placebo arm and apremilast arm; the null hypothesis is that the effects of the two treatment arms (i.e., placebo vs. apremilast) have no difference. Statistical tests will be at the two-sided 0.05 significance level and the corresponding p-values and two-sided 95% confidence intervals will be reported. For the primary and the major secondary endpoints, the sequential approach will be used for

multiplicity adjustment to control the overall type I error rate. The primary approach for treating missing values will be multiple imputation (MI) methods.

For binary endpoints, the two arms will be compared using Cochran–Mantel–Haenszel (CMH) test adjusted by stratification factors. For continuous endpoints, the analysis of covariance (ANCOVA) model will be performed. The ANCOVA model will have the change or percent change from baseline as the dependent variable and will include treatment arms and stratification factors as independent variables and the baseline value as a covariate variable.

9.2 Subject Accountability

The number and percent of subjects who were screened, randomized, received study treatment, and completed the study will be summarized by treatment group. The number and percent of subjects who discontinued study treatment, and study will be tabulated, along with the reason for discontinuation. The number of subjects included in and excluded from each analysis set and reason for exclusion will also be summarized.

Key study dates for the first subject enrolled, last subject enrolled, last subject's end of IP, and last subject's end of study will be presented.

9.3 Important Protocol Deviations

Important protocol deviations will be categorized by Clinical Research Physician (CRP). A list of important protocol deviations for all subjects for week 16 analysis will be defined prior to the database lock and unblinding (i.e., prior to the unblinding of double-blind placebo-controlled phase data). The list of important protocol deviations for all subjects for final analysis will be finalized prior to the final database lock. This listing of important protocol deviation for week 16 analysis will also identify which subjects are to be excluded from the per-protocol population.

Important protocol deviations will be summarized by treatment arm for Placebo-controlled Phase (Weeks 0 to16) and Apremilast Extension Phase (Weeks 16 to 52). Summary tables showing the number and percent of subjects with at least one important protocol deviations and by each category of important protocol deviations will be provided. Listings of subjects with important protocol deviations will be provided.

Additionally, protocol deviations and important protocol deviations related to COVID-19 will be summarized.

9.3.1 Treatment Compliance

As part of the routine recording of the amount of IP taken by each subject, the numbers of tablets dispensed and/or returned will be recorded at visits in treatment phases. These records will be used to calculate treatment compliance.

The treatment compliance for each subject will be computed as 100 times the total number of tablets taken (the total number of tablets dispensed minus the total number of tablets returned) over the analysis phase or period divided by the intended total number of tablets that should have been taken over the same phase or period.

Summary statistics for compliance (%) will be provided by treatment arm for each analysis phase or period. Frequency summary tables of compliance will also be presented with the following categories: < 75%, >= 75% to <= 120%, and > 120%. A subject data listing of drug accountability records will be provided.

9.3.1.1 Placebo-controlled Phase (Weeks 0 to 16)

Treatment compliance will be calculated for the treatment duration in the phase specified in Section 9.6.1 and will be summarized by treatment arm for ITT population.

9.3.1.2 Apremilast Extension Phase (Weeks 16 to 52)

Treatment compliance will be calculated for the treatment duration in the phase specified in Section 9.6.2 and will be summarized by treatment arm for the subset of ITT population who entered the phase.

9.3.1.3 Apremilast-exposure Period

Treatment compliance will be calculated for the treatment duration in the period specified in Section 9.6.3

Treatment compliance will be summarized for apremilast treatment for apremilast as treated population.

9.4 Subject Disposition

The number of subjects screened, the number and percentage of subjects randomized (as recorded in the IVRS database) will be summarized. The failed inclusion/exclusion criteria of subjects who were screened and not randomized will be included in the summary. The above percentages will be based on the number of subjects screened.

The number and percentage of subjects randomized will be tabulated by treatment arm, study site, country and region. The percentages will be based on the number of subjects randomized.

Subject disposition will be provided by treatment arm and phase:

- Placebo-controlled Phase – Weeks 0 to16

  Summary will be based on subjects who are randomized. Tabulation of subjects included in the ITT, PP, and Safety populations, subjects who completed and discontinued early will be provided. The primary reasons for early discontinuation will also be tabulated. Subjects who completed include subjects who have IP dispensed at Week 16/Visit 7, discontinued early subjects include those who discontinued prior to the Week 16 visit window.

- Apremilast Extension Phase – Weeks 16 to 52

  Summary will be based on subjects who entered the Apremilast Extension Phase, ie, who had IP dispensed for the Phase. Subjects who entered the phase, took at least one dose of IP, completed, discontinued early, and primary reason for discontinuation will be provided.

- Observational Follow-up Phase (14 weeks)

  Summary will include subjects who entered Follow-up Phase. Subjects who entered the phase, completed, discontinued early, and primary reason for discontinuation will be provided.

The primary reasons for discontinuation are collected in the screening disposition, treatment disposition and follow-up disposition eCRF using the following categories:

- Adverse Event
- Lack of efficacy
- Withdrawal by subject
- Withdrawal by parent/guardian
- Death
- Lost to follow-up
- Non-compliance with IP
- Protocol deviation
- Pregnancy
- Physician decision

- Study terminated by Sponsor
- Other

Discontinued reason due to COVID-19 controlled measure will also be provided.

9.5 Demographic and Baseline Characteristics

Summaries for the demographics, baseline characteristics, prior medication/procedure, and concomitant medications/procedure will be presented for the ITT population by treatment group and overall.

9.5.1 Demographics

Summary statistics will be provided for the following continuous variables:

- Age (years)
- Weight (kg)
- Height (cm)
- Baseline Body Mass Index (BMI; kg/m2)

Number and percentage will be provided for the following categorical variables:

- Sex (Male, Female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islanders, White, Other)
- Age category (6-11, 12-17 years)
- Baseline weight category (≥ 20 kg to < 50 kg, ≥ 50 kg)
- Age and baseline weight category (6-11 years and ≥ 20 kg to < 50 kg, 6-11 years and ≥ 50 kg, 12-17 years and ≥ 20 kg to < 50 kg, 12-17 years and ≥ 50 kg)
- Baseline BMI category (Underweight, Healthy Weight, Overweight, Obesity)
- Geographical region (USA, Canada, Europe, Rest of the World)

(Note: Subjects ≥ 20 kg to < 50 kg will receive apremilast 20 mg BID or placebo BID and subjects ≥ 50 kg will receive apremilast 30 mg BID or placebo BID. Baseline weight at randomization will be used to determine apremilast dose level for the whole study.)

9.5.2 Baseline disease characteristics

Baseline clinical characteristics will be summarized descriptively by treatment group, which will include the following:

- Duration of plaque psoriasis (from date of diagnosis to the date of informed consent; year, presented one digit after the decimal point)
- Duration of plaque psoriasis categories (< 1, ≥ 1 to < 2, ≥ 2 to <5, ≥ 5 years)
- Baseline sPGA score: Moderate (3), Severe (4)
- Baseline PASI score
- Baseline PASI score category
- Baseline BSA (%) score
- Baseline BSA (%) score category
- Baseline CDLQI total score
- Baseline CDLQI total score category: No effect on child's life (0-1); Small effect (2-6); Moderate effect (7-12); Very large effect (13-18); Extremely large effect (19-30)
- Baseline FDLQI total score
- Baseline ScPGA score
- Baseline sPGA-G score
- Baseline whole body itch NRS
- Baseline whole body itch NRS category (0 to 4, 5 to 10)
- Number of prior phototherapies
- Number of failed prior phototherapies
- Number of prior conventional systemic therapies
- Number of failed prior conventional systemic therapies
- Number of prior biologic therapies
- Number of failed prior biologic therapies
- Number of prior systemic therapies
- Number of failed prior systemic therapies

9.5.3 Medical History

Medical history terms will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or higher. A frequency summary (counts and percentage) of medical history will be presented by treatment group, system organ class (SOC), and preferred term (PT).

9.5.4 Concomitant Procedures

Concomitant procedures will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or higher. Frequency summaries of concomitant procedures will be provided for the safety population by treatment group, SOC, and PT.

A concomitant procedure is defined similarly as a TEAE. Concomitant procedures will be summarized similarly as TEAEs for (1) Placebo-controlled Phase (Weeks 0 to16), (2) Apremilast Extension Phase (Weeks 16 to 52), and (3) Observational Follow-up Phase (14 weeks).

9.5.5 Prior and Concomitant Medications

The Anatomical Therapeutic Chemical (ATC) coding scheme of the World Health Organization Drug Dictionary (WHO DD version March 2020 or later) will be used to group prior and concomitant medications and prior psoriasis medications into relevant categories. Frequency summaries will be provided by treatment group, ATC level, and standardized medication name for the safety population.

Prior medications are defined as those started before the first dose of IP (whether or not ended before the first dose of IP). Prior medications that continue after the first dose of IP will also be reported as concomitant medications.

Prior psoriasis medications will also be classified as phototherapies, biologic therapies, conventional systemic therapies, systemic therapies (including both biologic therapies and conventional systemic therapies).

For each of the two treatment phases, concomitant mediations are defined as non-study medications started during the phase, or started before the phase and ended or remain ongoing during the phase. The treatment phase for concomitant medications will start on the first dose date and end one day prior to the last dose date. For the Apremilast Exposure Period, concomitant mediations are similarly defined.

Medications in the Follow-up Phase will include those medications started on or after the date of last dose of investigational product.

Summaries will be provided for prior psoriasis medications and prior medications, as well as for concomitant medications in: (1) Placebo-controlled Phase (Weeks 0 to16), (2) Apremilast Extension Phase (Weeks 16 to 52) , (3) Apremilast Exposure Period, and (4) Observational Follow-up Phase (14 weeks).

The following two subsets of concomitant medications will also be listed separately and summarized for both treatment phases and the Apremilast Exposure Period:

- Moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications
- Low potency topical corticosteroids

9.6 Treatment Duration

Treatment duration will be summarized by treatment group for the analysis phases and for the Apremilast-exposure Period. Subjects who are treated in the corresponding phases or period will be used, i.e., the safety population for Placebo-controlled Phase (Weeks 0 to16), subjects who are treated in Apremilast Extension Phase (Weeks 16-52), and the apremilast subjects as treated population for the Apremilast-exposure Period.

Summary statistics for treatment duration (in weeks), as well as a frequency summary of treatment duration categories (e.g., < 4, ≥ 4 - < 8 weeks, etc.), will be provided.

9.6.1 Placebo-controlled Phase (Weeks 0 to 16)

Treatment duration (in weeks) is calculated from the date of the first dose of IP at Week 0/Visit 2 to either the date one day prior to the first dose date in the Apremilast Extension phase for the IP dispense at Week 16/Visit 7, or the date of the last dose of IP in the study for subjects who discontinue in the phase. Imputation rule for partially or completely missing last dose date is specified in Appendix C.

Treatment duration will be summarized by actual treatment (placebo or apremilast) for the safety population.

9.6.2 Apremilast Extension Phase (Weeks 16 to 52)

Treatment duration is calculated from the first dose date in the Apremilast Extension phase for the IP dispense at Week 16/Visit 7 to the date of the last dose of IP in the study for subjects who discontinue in the phase or who complete the study at Week 52/Visit 16. Imputation rule for partially or completely missing last dose date is specified in Appendix C.

9.6.3 Apremilast-exposure Period

Treatment duration for Apremilast-exposure Period is calculated from the date of the first dose of apremilast, which is the date of the first dose of apremilast after randomization at Week 0/Visit 2 or switched to apremilast at Week 16/Visit 7, to the last apremilast dose date for subjects who discontinue in the first 52 weeks or who complete the study at Week 52/Visit 16. Imputation rule for partially or completely missing last dose date is specified in Appendix C.

9.7 Efficacy Analyses

9.7.1 Multiplicity Adjustment

The primary and the major secondary efficacy endpoints will be hierarchically ranked for testing to control the overall type I error rate in claiming statistical significance at the two-sided 0.05 significance level. Specifically, for the primary efficacy endpoint (sPGA response at Week 16), if the two-sided p-value from the comparison between apremilast arm and placebo arm is below 0.05, the outcome will be considered statistically significant and apremilast will be declared effective. For the major secondary endpoint, statistical significance will be claimed only if its two-sided p-value is below 0.05 and tests for the primary endpoint is significant at the two-sided 0.05 level. The proposed test sequence for the primary and the major secondary efficacy endpoints is listed as the following:

- Proportions of subjects who achieve sPGA response at Week 16 (defined as sPGA score of clear (0) or almost clear (1) with at least a 2-point reduction from baseline at Week 16)
- Proportion of subjects who achieve PASI-75 response at Week 16 (defined as at least a 75% reduction in total PASI score from baseline at Week 16)

For the other secondary efficacy endpoints, multiplicity adjustment will not be applied and statistical significance will not be claimed.

9.7.2 Analyses of Primary Efficacy Endpoint(s)/Estimand(s)

The primary endpoint is the proportions of subjects who achieving sPGA response at Week 16 (defined as sPGA score of clear (0) or almost clear (1) with at least a 2-point reduction from baseline at Week 16). The estimand of primary interest is defined in terms of the following four attributes:

A. The population is defined through appropriate inclusion/exclusion criteria to reflect the targeted patient population for approval (e.g., ITT Population);

B. The variable is the proportion of subjects who achieving sPGA response at Week 16. Subjects who terminate early due to lack of efficacy or add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16 will be assigned as non-responders.

C. Potential intercurrent events are captured through the variable definition (B).

D. The summary measure is the difference in response proportions.

Missing values at Week 16 for this estimand will be imputed using the multiple imputation (MI) method (SAS Institute Inc. 2011) based on similar subjects who remained in the study as the primary method. With this missing data handling approach, the estimand answers a clinically relevant question that compares the number of subjects who achieve the defined response criteria at Week 16 in the ITT population.

The primary endpoint will be analyzed using the CMH (Cochran–Mantel–Haenszel) test adjusting for the stratification factor at randomization. The two-sided p-values from the CMH test, the adjusted treatment difference in proportion using the weighted average of the treatment differences across the strata with the CMH weights, along with the associated two-sided 95% confidence intervals (CIs) using a normal approximation to the weighted average will be provided.

Sensitivity analysis will be conducted to account for missing data using the last observation carried forward (LOCF) method, the non-responder imputation (NRI) method and the tipping point analyses. A sensitivity analysis using the PP population will also be performed.

For the multiple imputation method, the SAS procedure MI will be used to impute missing sPGA scores at the scheduled analysis visits in the Placebo-controlled Phase (Weeks 0-16) to create M=25 complete data sets. The sPGA assessment data will not be included in multiple imputation steps from subjects who terminate early due to lack of efficacy or added moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16. The missing data patterns will be checked at the scheduled analysis visits, ie, Baseline (Week 0), and Weeks 2, 4, 8, 12 and 16. If there are non-monotone missing patterns, two separate imputation procedures will be used to complete the imputation process.

In the first step, the Markov Chain Monte Carlo (MCMC) method will be used with a single chain to impute missing scores by treatment and stratification factor to create M=25 imputed data sets with monotone missing patterns. In case there are

convergence issues, a simple model will be used to impute the missing scores by treatment, with further simplification by dropping both treatment and stratification factor in imputation model if necessary. The seed will be set to 17813721. The imputed scores will be rounded to the nearest integer. The minimum and the maximum values for imputation will be 0 and 4, which correspond to the lowest and the highest sPGA scores.

In the second step, the predictive mean matching method will be used to impute the remaining missing scores for the 25 data sets with monotone missing patterns. The imputation procedure will use monotone statement to create one complete data set for each of the monotone data set from the first step, and the variables will include treatment arm, stratification factor, and sPGA scores at scheduled analysis visits from baseline to Week 16. The seed will be set to 55218163.

After the completion of imputation, sPGA response at Week 16 will be derived based on both observed and imputed scores. Data from subjects who terminate early due to lack of efficacy or add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16 as non-responders will be added to the 25 imputed data sets. The same CMH method will be used to analyze the 25 complete data sets and the SAS procedure MIANALYZE will be used to combine the results for the statistical inferences.

The primary endpoint will be assessed for site effect as specified in Section 9.7.6.

9.7.3 Analyses of Major Secondary Efficacy Endpoint(s)

The major secondary efficacy endpoint is the proportions of subjects who achieve PASI-75 response at Week 16 (defined as at least a 75% reduction in total PASI score from baseline at Week 16), which will be analyzed using the same approaches as the primary endpoint. The estimand of primary interest is defined in terms of the following four attributes:

A. The population is defined through appropriate inclusion/exclusion criteria to reflect the targeted patient population for approval (e.g., ITT Population);

B. The variable is the proportion of subjects who achieving PASI-75 response at Week 16. Subjects who terminate early due to lack of efficacy or add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16 will be assigned as non-responders.

C. Potential intercurrent events are captured through the variable definition (B).

D. The summary measure is the difference in response proportions.

The primary analysis for this estimand will use CMH (Cochran–Mantel–Haenszel) test adjusting for the stratification factor at randomization. Missing values at Week 16 will be imputed using the multiple imputation (MI) method (SAS Institute Inc. 2011) based on similar subjects who remained in the study. The two-sided p-values from the CMH test, the adjusted treatment difference in proportion using the weighted average of the treatment differences across the strata with the CMH weights, along with the associated two-sided 95% CIs using a normal approximation to the weighted average will be provided.

Sensitivity analysis will be conducted to account for missing data using the last observation carried forward (LOCF) method, the non-responder imputation (NRI) method, and tipping point analyses. A sensitivity analysis using the PP population will also be performed.

For using the multiple imputation method, the SAS procedure MI will be used to impute missing total PASI scores at the scheduled analysis visits in the Placebo-controlled Phase (Weeks 0-16) to create M=25 complete data sets. The PASI assessment data will not be included in multiple imputation steps from subjects who terminate early due to lack of efficacy or who add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16.

The missing data patterns will be checked at the scheduled analysis visits, ie, Baseline (Week 0), and Weeks 2, 4, 8, 12 and 16. If there are non-monotone missing patterns, two separate imputation steps will be used to complete the imputation process.

In the first step, the Markov Chain Monte Carlo (MCMC) method will be used to impute missing scores by treatment and stratification factor to create M=25 imputed data sets with monotone missing patterns. In case there are convergence issues, a simple model will be used to impute the missing scores by treatment, with further simplification by dropping both treatment and stratification factor in imputation model if necessary. The minimum and the maximum values for imputation will be 0 and 72, which correspond to the lowest and the highest total PASI scores. The seed will be set to 17813721 and a single chain will be used to produce imputations. The imputed scores will be rounded to numbers with one digit (format xx.x).

In the second step, the predictive mean matching method will be used to impute the remaining missing scores for the 25 data sets with monotone missing patterns. The

imputation procedure will use monotone statement to create one complete data set for each of the monotone data set from the first step, and the variables will include treatment arm, stratification factor, and total PASI scores at scheduled analysis visits from baseline to Week 16. The seed will be set to 55218163.

After the completion of imputation, PASI-75 response at Week 16 will be derived based on both observed and imputed scores. Data from subjects who terminate early due to lack of efficacy or add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16 as non-responders will be added to the 25 imputed data sets. The same CMH method will be used to analyze the 25 complete data sets and the SAS procedure MIANALYZE will be used to combine the results for the statistical inferences.

The major secondary endpoint will be assessed for site effect as specified in Section 9.7.6.

9.7.4 Analyses of Other Secondary Efficacy Endpoint(s)

The other secondary efficacy endpoints will be analyzed similarly as the primary and the major secondary endpoints. The main analysis will be based on the ITT population with missing values imputed using multiple imputation methods. The two-sided p-values and two-sided 95% confidence intervals (CIs) will be reported for treatment difference between placebo arm and apremilast arm. Multiplicity adjustment will not be applied and statistical significance will not be claimed.

9.7.4.1 Binary Variables

The binary variables for secondary endpoints include the following:

- Proportion of subjects who achieve PASI-50 response at Week 16 (defined as at least a 50% reduction in total PASI score from baseline at Week 16)
- Proportion of subjects who achieve CDLQI (0/1) response at Week 16 (defined as CDLQI score 0 or 1 at Week 16) for subjects with baseline CDLQI score 2 or above

For these two binary endpoints, the treatment difference between apremilast arm and placebo arm will be compared using CMH (Cochran–Mantel–Haenszel) test adjusting for the stratification factor at randomization. The two-sided p-values from the CMH test, the adjusted treatment difference in proportion using the weighted average of the treatment differences across the strata with the CMH weights, along with the associated two-sided 95% CIs using a normal approximation to the weighted average will be provided.

Subjects who terminate early due to lack of efficacy or add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16 will be considered as non-responders. Missing values will be imputed using the similar MI method as the primary and major secondary endpoints, with sensitivity analysis using the LOCF method and NRI method.

9.7.4.2 Continuous Variables

The continuous variables for secondary endpoints include the following:

- Percent change from baseline in total PASI score at Week 16
- Percent change from baseline in affected BSA at Week 16
- Change from baseline in CDLQI score at Week 16

The three continuous endpoints will be analyzed based on the ITT population using the analysis of covariance (ANCOVA) model. The ANCOVA model will use the change or percent change from baseline as the dependent variable and will include treatment group and stratification factor as independent variables and the baseline value as a covariate variable. Within-group least-squares (LS) mean changes from baseline at Week 16, the associated standard errors (SEs) and two-sided 95% CIs, treatment differences in LS mean changes from baseline, and the associated two-sided 95% CIs and p-values, will be derived from the ANCOVA model. Missing values at Week 16 will be imputed using the MI method, with sensitivity analysis using the LOCF method. The primary analysis will estimate treatment effect without additional psoriasis medications, ie, data from assessments after adding those medications will be removed before using MI method for subjects who add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16. A sensitivity analysis will be performed with all available data using MI method regardless whether subjects who add moderate-to-high potency topical steroid preparations (early escape) or other protocol prohibited medications prior to Week 16.

9.7.5 Analyses of Exploratory Efficacy Endpoint(s)

Descriptive statistics will be provided for all endpoints by study phases. Data summaries will be provided for scheduled timepoints in each phase; and for an end of phase assessment in each treatment phase (the last assessment for each subject within the phase). Specifically, for continuous variables, descriptive statistics for assessment value and change or percent changes from baseline will be provided. Categorical variables will

be summarized with frequency tabulations for specified time points. In addition, two-sided 95% CIs will be provided for each treatment.

For the exploratory endpoints evaluated at Week 16, statistical comparisons between the two treatment arms (apremilast versus placebo) will be performed using the ITT population when appropriate. No multiplicity adjustment will be pre-specified and nominal two-sided p-values will be presented. Missing values at Week 16 will be imputed using the LOCF method as the main approach. For binary endpoints, the treatment difference will be compared using CMH test adjusting for the stratification factor at randomization. For continuous endpoints, the analyses will use an analysis of covariance (ANCOVA) model with the change or percent change from baseline as the dependent variable and will include treatment group and stratification factor as independent variables and the baseline value as a covariate variable.

9.7.5.1 Study Phases

The subjects included and the treatment arms for statistical analyses and summaries in each phase are described below in this section. The specific summaries are discussed in the following sections.

- Placebo-controlled Phase (Weeks 0 to 16)

  Subjects in the ITT population will be put in the two treatment arms for analysis:

  - Placebo
  - Apremilast

- Apremilast Extension Phase (Weeks 16 to 52)

  Subjects who entered the extension phase will be included in analysis, ie, who were initially randomized to apremilast and continued at Week 16 and who were initially randomized to placebo and switched to apremilast at Week 16. The two treatment arms are:

  - Placebo/Apremilast
  - Apremilast/Apremilast

- Observational Follow-up Phase (14 weeks)

  For subjects who entered Follow-up Phase from Placebo-controlled Phase, summaries will be provided by their initial treatment arm:

  - Placebo
  - Apremilast

For subjects who entered Follow-up Phase from Apremilast Extension Phase, summaries will be provided by their treatment sequence:

- Placebo/Apremilast
- Apremilast/Apremilast

9.7.5.2 sPGA, PASI and BSA

The endpoints include the following:

- Proportion of subjects who achieve PASI-75, PASI-50, and PASI-90 response at all visits
- Proportion of subjects who achieve sPGA (0/1) response at all visits
- Percent changes from the baseline in total PASI score at all visits
- Percent changes from the baseline in affected BSA at all visits

All endpoints will be summarized by time point (ie, visit) using descriptive statistics. PASI-90 response at Week 16 will also be analyzed using CMH test.

9.7.5.3 ScPGA and Modified sPGA-G

The endpoints include:

- ScPGA: Proportion of subjects with scalp psoriasis with improvement of ScPGA of clear (0) or almost clear (1)
- Modified sPGA-G: Proportion of subjects with genital psoriasis with improvement of sPGA-G of clear (0) or almost clear (1)

The endpoints will be summarized by time point using descriptive statistics. Week 16 proportions between the two arms will be compared using CMH test.

In addition, shift tables will be provided from baseline to post-baseline visits for the corresponding full range of scores.

9.7.5.4 CDLQI and FDLQI

The endpoints include:

- CDLQI (0/1): Proportion of subjects who achieve CDLQI (0/1)
- CDLQI: Change from baseline in CDLQI total score
- FDLQI (0/1): Proportion of subjects achieving FDLQI (0/1)
- FDLQI: Change from baseline in FDLQI total score

All endpoints will be summarized by time point using descriptive statistics. CDLQI (0/1) response at Week 16 will also be analyzed using CMH test. Change from baseline for CDLQI total score at Week 16 will be analyzed using ANCOVA model.

9.7.5.5 Modified Whole Body Itch Numeric Rating Scale (NRS)

The endpoints include:

- NRS response: Proportion of subjects with ≥4-point reduction (improvement) from baseline in whole body itch NRS score
- Change from baseline in whole body itch NRS score

The two endpoints will be summarized by time point using descriptive statistics. NRS response at Week 16 will also be analyzed using CMH test. Change from baseline in whole body itch NRS score at Week 16 will be analyzed using ANCOVA model.

9.7.6 Assessing Study Site Effect

This study is a multicenter study and planned to have more than 100 study sites to enroll at least 230 subjects. A single site may not have sufficient number of subjects to allow meaningful within-site analysis; therefore, study sites will be pooled for the analysis. The primary and the major secondary efficacy endpoints will be further analyzed adjusting for the stratification factor and the pooled site and examining whether the treatment differences are consistent with those from the primary analysis.

In addition, the consistency of the treatment effect across individual study sites (or pooled sites) will also be assessed by performing a subgroup-type analysis with individual study sites (or pooled sites) treated as subgroups. Listings of response rates will be provided by individual study site and by pooled site. The treatment difference for each of the individual study sites (or pooled sites) will be reviewed to evaluate the effect among the individual study sites (or pooled sites).

In pooling sites for analysis, the minimum cell size of 5 randomized subjects will be used. With five subjects per treatment arm (placebo arm or apremilast arm) per stratum of age group (6-11 or 12-17 years old), there is an approximately 67% chance to observe at least one sPGA responder per stratum for the apremilast arm (assuming a 20% responder rate) and 23% chance to observe at least one sPGA responder per stratum for the placebo arm (assuming a 5% responder rate).

Site will be pooled within the 4 geographical regions (USA, Canada, Europe, Rest of the World). In each region, the smallest sites will be pooled first until the pooled site has a

minimum cell size of 5 for all four cells. The remaining unpooled sites will be pooled with the smallest pooled site within the region.

The 4 geographical regions will be ranked from the smallest to the largest according to the total number of randomized subjects. If the smallest region can't reach a minimum cell size of 5 for all four cells, all sites in that region will be pooled with the smallest pooled site in the next region.

In analysis using CMH method, if a pooled site does not have a responder in one of the age groups (6-11 or 12-17 years old), that pooled site will be further pooled with the next pooled site in the region.

9.8 Safety Analyses

Safety will be assessed via descriptive statistics and point estimates. Unless otherwise specified, all safety analyses described in this section will be performed for both the Placebo-controlled Phase and the Apremilast-exposure Period. The safety analyses for the Placebo-controlled Phase will be based on the safety population and presented by treatment group (placebo arm or apremilast arm).The safety analyses for the Apremilast-exposure Period will be based on the apremilast subjects as treated population and presented for both groups (subjects who were initially treated with placebo and switched to apremilast at Week 16, or subjects who were initially treated with apremilast from Week 0), and for overall apremilast group irrespective of the start time of apremilast exposure (at Week 0 or 16). For apremilast treatment groups, the further refined summaries for both dose levels (20 mg BID or 30 mg BID) will also be provided.

For the analyses of AEs and marked abnormalities, the following point estimates are distinguished:

- Subject incidence: Subject incidence (i.e., percentage [%] used in a frequency summary) is defined as the number of subjects with the specific event divided by the number of subjects included in the analysis. Subjects with multiple occurrences of the specific event in the specific analysis period will be counted only once in the numerator.
- Exposure-adjusted incidence rate (EAIR) per 100 subject-years: The EAIR per 100 subject-years is defined as 100 times the number of subjects with the specific event divided by the total exposure time (in years) among subjects included in the analysis. Subjects with multiple occurrences of the specific event in the specific analysis period will be counted only once in the numerator. The

exposure time for a subject without the specific event is the treatment duration, whereas the exposure time for a subject with the specific event is the treatment duration up to the start date (inclusive) of the first occurrence of the specific event. For subject only having the specific event started after and within the 28 days of the last dose date, the exposure time will be calculated to the last dose date. The total exposure time in years is calculated by dividing the sum of exposure time in days over all subjects included in the analysis by 365.25. The EAIR per 100 subject-years is interpreted as the expected number of subjects with at least one occurrence of the specific event per 100 subject-years of exposure to the IP.

AEs and marked abnormalities will be summarized by subject incidence and EAIR for the Placebo-controlled Phase (Weeks 0 to16) and for the Apremilast-exposure Period. In addition, selected summaries for the first 16 weeks in the Apremilast Extension Phase (Weeks 16 to 52) and for the Observational Follow-up Phase (14 weeks) after 28 days of the last dose date will be presented. Selected summaries will also be presented for the first 16 weeks in Apremilast-exposure Period for apremilast group irrespective of the start time of apremilast exposure (at Week 0 or 16).

Descriptive statistics will be provided for vital signs, weight, height and BMI, tanner staging of sexual development, stool diary for diarrhea, Columbia-Suicide Severity Rating Scale (C-SSRS), laboratory values (continuous measurements) by treatment and visit, including the end of treatment visits. The baseline value, value at the time point, and change from baseline will be summarized for subjects who have values at baseline and at the time point.

Shift tables, that is, tables that summarize the baseline categories (normal, abnormal) versus the category at the end of the respective periods or versus the worst post-baseline category, include subjects who have values at baseline and at least one post-baseline value. Similarly, in frequency summaries of shifts from baseline at scheduled study weeks per protocol, only subjects who have values at baseline and at the time point will be included.

9.8.1 Adverse Events

AEs will be coded according to the MedDRA version 24.0 or higher. Unless otherwise specified, AEs will be summarized by SOC and PT, with SOCs presented in the standard international order and PTs within SOCs will be presented in descending order of subject incidence.

9.8.1.1 Overall Summary of TEAEs

An overall summary of the following TEAE categories will be provided for the Placebo-controlled Phase (Weeks 0 to16) for subjects with:

- Any TEAE
- Any drug-related TEAE
- Any severe TEAE
- Any serious TEAE
- Any serious drug-related TEAE
- Any TEAE leading to drug interruption
- Any TEAE leading to drug withdrawal
- Any TEAE leading to death

In addition, for Apremilast Extension Phase (Weeks 16 to 52), Observational Follow-up Phase (14 weeks) and Apremilast Exposure Period overall summary of TEAEs will also be provided.

9.8.1.2 All TEAEs

All TEAEs will be summarized by SOC and PT as well as by PT only (in descending order of subject incidence). In addition to the Placebo-controlled Phase (Weeks 0 to16) and the Apremilast-exposure Period, summaries will also be provided for the Apremilast Extension Phase (Weeks 16 to 52).

New events of all TEAEs by exposure interval (≤ 1, > 1 to ≤ 4, > 4 to ≤ 8, > 8 to ≤ 12, and > 12 Weeks) will be summarized for the Placebo-controlled Phase. Each subject is counted once for either subject incidence or EAIR for each applicable specific TEAE in each exposure interval where an event started. The denominator of a subject incidence is the number of subjects with treatment duration exceeding the lower bound of the particular exposure interval, while the denominator of an EAIR is the sum of the exposure time during the exposure interval (up to the first event start date for subjects with at least one event starting in the interval) among the same number of subjects as in the denominator of the corresponding subject incidence.

In addition, new events of all TEAEs by exposure interval (≤ 1, > 1 to ≤ 8, > 8 to ≤ 16, > 16 to ≤ 24, > 24 to ≤ 32, > 32 to ≤ 40, > 40 to ≤ 48, > 48 Weeks) will be summarized for the Apremilast-exposure Period.

All TEAEs will be summarized by age category (6-11, 12-17 years), sex, and race. All TEAEs will also be summarized by patients using or not using low potency topical corticosteroids.

All TEAEs occurring after the date of the last dose of IP will also be summarized by SOC and PT for the Observational Follow-up Phase (14 weeks). The EAIR will not be provided in this summary.

9.8.1.3 Common TEAEs

TEAEs with subject incidence ≥ 5% (or another cut-off if justified) in any treatment group will be summarized by SOC and PT as well as by PT only in descending order of subject incidence.

9.8.1.4 Drug-related TEAEs

Drug-related TEAEs will be summarized and new events of drug-related TEAEs by exposure interval (see Section 9.8.1.2) will be summarized.

9.8.1.5 TEAEs by Maximum Severity

All TEAEs will be summarized by maximum severity (mild, moderate, severe, and, if needed, missing). If a subject reports multiple occurrences of a specific event within a specific analysis phase or period, the subject will be counted only once by the maximum severity. If the severity is missing for one or more of the occurrences, the maximum severity of the remaining occurrences will be used. If the severity is missing for all of the occurrences, the subject will be counted only once in the “missing” category of severity.

9.8.1.6 Serious TEAEs

Serious TEAEs and serious drug-related TEAEs will be summarized.

New events of serious TEAEs and serious drug-related TEAEs by exposure interval will be summarized for Placebo-controlled Phase (Weeks 0 to16).

Serious TEAEs will be summarized by age category (6-11, 12-17 years), sex.

A subject data listing of all serious AEs (both TEAEs and non-TEAEs) will be provided.

9.8.1.7 TEAEs Leading to Drug Interruption and TEAEs Leading to Drug Withdrawal

TEAEs leading to drug interruption and TEAEs leading to drug withdrawal will be summarized.

TEAEs leading to drug withdrawal will also be summarized by age category (6-11, 12-17 years), sex, and patients using or not using low potency topical corticosteroids.

A subject data listing of TEAEs leading to drug withdrawal will be provided.

9.8.1.8 Deaths

TEAEs leading to death will be summarized. A subject data listing of all deaths will be provided.

9.8.1.9 Onset and duration

For selected TEAE events, categories for onset (start) day from the first dose date and event duration (from start date to end date) will be summarized.

9.8.2 Laboratory Test Results

Summary statistics of observed values and changes from baseline in laboratory parameters will be provided over time. Frequency summaries (shift tables) of shifts from baseline to post-baseline time points and to the worst post-baseline value in terms of normal/abnormal will be provided.

Laboratory marked abnormalities will be summarized; subject incidence and EAIR for each abnormality will be calculated based on subjects with a baseline value and at least one post-baseline value for criteria requiring baseline or subjects with at least one post-baseline value for criteria not requiring baseline. A subject data listing of laboratory marked abnormalities will be provided. Laboratory marked abnormalities will also be summarized for subjects with normal values at baseline and for subjects with abnormal values at baseline separately. For the purposes of these summaries, subjects with abnormal (normal) values at baseline are defined as those whose baseline value is low (not low) for criteria concerning low values and those whose baseline value is high (not high) for criteria concerning high values; both low and high are relative to the laboratory normal range.

Subject data listings for clinical laboratory data, including hematology, serum chemistry and urinalysis, will be provided.

9.8.3 Vital Signs

Summary statistics of observed values and changes from baseline in vital signs will be provided over time. Frequency summaries (shift tables) of shifts from baseline to post-baseline time points and to the worst post-baseline value in terms of normal/abnormal will be provided for pulse and blood pressure.

A subject data listing of all vital signs will be provided.

9.8.4 Physical Measurements

The protocol specified that physical examination findings of clinical significance (as defined by the investigator) are to be reported as AEs. No summary of physical examination findings will be provided.

9.8.5 Electrocardiogram

Not applicable.

9.8.6 Antibody Formation

Not applicable.

9.8.7 Exposure to Concomitant Medication

The Anatomical Therapeutic Chemical (ATC) coding scheme of the World Health Organization Drug Dictionary (WHO DD version March 2020 or later) will be used to group prior and concomitant medications and prior psoriasis medications into relevant categories. Frequency summaries will be provided by treatment group, ATC level, and standardized medication name for the safety population.

Summaries will be provided for prior psoriasis medications and prior medications, as well as for concomitant medications in: (1) Placebo-controlled Phase (Weeks 0 to16), (2) Apremilast Extension Phase (Weeks 16 to 52), (3) Apremilast Exposure Period, and (4) Observational Follow-up Phase (14 weeks).

9.8.8 Exposure to Concomitant Procedures

Concomitant procedures will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or higher. Frequency summaries of concomitant procedures will be provided for the safety population by treatment group, SOC, and PT.

Concomitant procedures will be summarized similarly as concomitant medication for (1) Placebo-controlled Phase (Weeks 0 to16), (2) Apremilast Extension Phase (Weeks 16 to 52), (3) Apremilast Exposure Period, and (4) Observational Follow-up Phase (14 weeks).

9.8.9 Stool Diary for Diarrhea

Subjects and their parent/guardian will be supplied with paper diaries that will be filled out daily to record and describe any diarrhea, including duration, frequency, treatment, and associated symptoms. Subjects and their parent/guardian will complete the Stool Diary every day beginning after the first dose of apremilast, up to the 4-week post last dose follow-up visit.

Summary and subject data listing of stool diary entries will be provided.

9.8.10 Psychiatric Evaluation

Subjects will complete the Columbia-Suicide Severity Rating Scale (C-SSRS) assessment. This questionnaire is suitable for assessment of suicidal ideation and behavior in clinical and research settings. The assessment will be completed at screening, baseline, and at post-baseline visits.

Summary and subject data listing of psychiatric evaluation results will be provided.

9.8.11 Tanner Staging of Sexual Development

Assessments of sexual maturity (Tanner Staging) will be performed at baseline and at the end of treatment for boys and girls. Summary and subject data listing of Tanner Staging assessment findings will be provided.

9.8.12 Psoriasis Flare and Rebound

Subjects with psoriasis flare or rebound will be summarized for the two treatment phases and for the Apremilast-Exposure Period. Subject data listing will be presented for psoriasis adverse events.

9.8.13 Weight, Height and BMI

Body weight and height will be measured by study visit. BMI will be calculated using weight and height.

Body weight, height and BMI will be summarized using change and percent change from baseline. Weight change (kg) and percent change (%) from baseline will also be summarized using categories as < -20, ≥ -20 to < -10, ≥-10 to < -5, ≥-5 to < 0, 0, > 0 to ≤ 5, > 5 to ≤ 10, > 10 to ≤ 20, and > 20 at the end of treatment phase or period.

The categories of weight change (kg) and percent change (%) are < -20, ≥ -20 to < -10, ≥ -10 to < -5, ≥- 5 to < 0, 0, > 0 to ≤ 5, > 5 to ≤ 10, > 10 to ≤ 20, and > 20. The categories of weight change (kg) and percent change (%) are < -20, ≥ -20 to < -10, ≥ -10 to < -5, ≥ -5 to < 0, 0, > 0 to ≤ 5, > 5 to ≤ 10, > 10 to ≤ 20, and > 20.BMI will also be summarized for shifting from baseline to post-baseline timepoints using the four categories as underweight, healthy weight, overweight or obesity.

Data listing will be provided for weight, height and BMI. Subjects with weight decrease from baseline ≥ 10% will be listed separately.

9.9 Other Analyses

9.9.1 Analyses of COVID-19 impacted data

Analyses will be performed to assess the impact of COVID-19 on the amount of data collected for the study.

[redacted]

10. Changes From Protocol-specified Analyses

There are no changes to the protocol-specified analyses.

11. Literature Citations / References

Basra MKA, Sue-Ho R, Finlay AY. The Family Dermatology Life Quality Index; measuring the secondary impact of skin disease. Br J Dermatol 2007; 156: 528-538. Erratum: British Journal of Dermatology 2007; 156: 791 (Correct FDLQI given here).

Basra M K A, Finlay A Y. The family impact of skin diseases: the greater patient concept. Br J Dermatol 2007; 156: 929-937.

Cole TJ, Green PJ (1992). Smoothing reference centile curves: the LMS method and penalized likelihood. Statistics in Medicine, 11:1305–1319.

Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index (CDLQI): Initial validation and practical use. British Journal of Dermatology, 1995; 132: 942-949.

Mehrotra, DV, and Railkar, R. Minimum risk weights for comparing treatments in stratified binomial trials. Statist. Med., 2000; 19: 811–825.

MKA Basra, O Edmunds, MS Salek, AY Finlay. Measurement of family impact of skin disease: further validation of the Family Dermatology Life Quality Index (FDLQI). J Eur Acad Dermatol Venereol 2008. 22: 813-821.

Naegeli, AN, Flood, E, Tucker J, Devlen J, Edson-Heredia E. The Worst Itch Numeric Rating Scale for patients with moderate to severe plaque psoriasis or psoriatic arthritis. Int J Dermatol 2015; 54:715-22.

Rigby RA, Stasinopoulos DM (2004). Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. Statistics in Medicine, 23:3053–3076.

SAS Institute Inc. 2011. SAS/STAT® 9.3 User's Guide. Cary, NC: SAS Institute Inc.

Centers for Disease Control (CDC) National Center for Health Statistics Clinical Growth Charts, Modified 2000: https://www.cdc.gov/growthcharts/clinical_charts.htm.

Kuczmarski RJ, Ogden CL, Guo SS, et al. 2000 CDC growth charts for the United States: Methods and development. National Center for Health Statistics.Vital Health Stat 11(246). 2002.

12. Prioritization of Analyses

Not applicable.

13. Data Not Covered by This Plan

[redacted]

14. Appendices

Appendix A. Criteria for Marked Abnormalities Value

Category / Analyte	**SI Units**	Criteria
Chemistry/ Alanine Aminotransferase (SGPT)	U/L	> 2*ULN
Albumin	Kg/m3	< 25
Alkaline Phosphatase	U/L	> 400
Aspartate Aminotransferase (SGOT)	U/L	> 2*ULN
Total Bilirubin	µmol/L	> 2*ULN
Total Bilirubin and Alanine Aminotransferase /Aspartate Aminotransferase	µmol/L and U/L	Bilirubin Value > 2xULN with (ALT or AST value > 2xULN)
Blood Urea Nitrogen	mmol/L	> 24
Calcium	mmol/L	< 1.8 > 3.0
Cholesterol	mmol/L	> 7.8
Creatinine	µmol/L	> 1.5*ULN
Glucose	mmol/L	< 2.8 > 13.9
Hemoglobin A1C	%	> 6.5
Lactate Dehydrogenase (LDH)	U/L	> 2*ULN
Magnesium	mmol/L	> 1.2
Phosphate	mmol/L	<1.03 >1.94
Potassium	mmol/L	<3.0 >5.4
Sodium	mmol/L	<132 >147
Triglycerides	mmol/L	> 3.4
Urate	umol/L	Male: > 480 Female: > 480
Hematology/ Hemoglobin	g/L	Female <110, Male <110 Female >150, Male >150
Leukocytes	10^9/L	< 2.0
Lymphocytes	10^9/L	< 1.0
Neutrophils	10^9/L	< 1.5

Category / Analyte	**SI Units**	Criteria
Platelets	10^9/L	<100 >500

Appendix B. Analytical Windows

Post baseline time points in all analyses will be captured based on analysis visit window (range of study days) around the target day for each analysis visit, based on the actual day of evaluation relative to a reference date. Appropriate dates will be used to calculate the study day, e.g., date of measurement or date of specimen collection will first be used, and then the date of visits/study weeks as recorded in the database will be used. If there are multiple measurements within a time point based on the study day, then the non-missing value from the closest measurement to the planned study day will be used for that visit. If the value at a scheduled visit is missing and there is no value available within the time window based on the study day, the value at the study week will be missing.

Table 1. Table for Visit Mapping for by Time Point Analysis

Analysis visit	Target Day	Visit Window
Placebo-controlled Phase		
Week 0 (Baseline)	1	<= 1
Week 2	15	2 to 21
Week 4	29	22 to 42
Week 8	57	43 to 70
Week 12	85	71 to 98
Week 16	113	99 to End of Phase
Apremilast Extension Phase		
Week 20	141	(End of Previous Phase + 1) to 154
Week 24	169	155 to 182
Week 28	197	183 to 210
Week 32	225	211 to 238
Week 36	253	239 to 266
Week 40	281	267 to 294
Week 44	309	295 to 322
Week 48	337	323 to 350
Week 52	365	351 to End of Phase

Note: Target day and visit window are relative to the date of Visit 2/Week 0 (Day 1) randomization date. For efficacy and safety analysis, definitions for start or end dates for a phase or period are specified in Section 5.3.

Time points in the analyses or summaries of efficacy data over time include the scheduled study weeks per protocol, the end of a study phase, and the observational follow-up visits. Appropriate dates (e.g., date of measurement or date of specimen

collection) will first be used to ensure only data (including data from scheduled, unscheduled, discontinuation, and observational follow-up visits) measured or collected within the specific analysis phase being analyzed or summarized are included, and then the visits/study weeks as recorded in the database will be used to assign one value (possibly missing) to each time point at the subject level. It is possible that multiple assessment values will fall into the same visit window. The following rule may be used to select the unique value for that analysis visit:

1. Among all assessments in the same visit window for the analysis visit, select the value with the assessment date closest to the target day of the analysis visit;

2. If the relative days from 2 assessments are equally close to, but on different sides of the target day, then the latter assessment will be used for that analysis visit;

3. If multiple assessments are available on the same relative day, then the highest value of these assessments will be used for that relative day.

For the Apremilast-exposure Period, the scheduled study weeks for placebo subjects who are treated with apremilast 20 or 30 mg BID at Week 16/Visit 7 will be mapped to reflect the study weeks relative to the first dose of apremilast for summaries of safety data (laboratory parameters, vital signs, etc).

Table 2. Adjustment of Study Weeks for Placebo Subjects Who Are Treated with Apremilast 20 or 30 mg BID after Week 16 in Summary of Safety Data over Time

Original Visit	Re-Mapped Visit for subjects initially randomized to placebo
Week 16	Week 0 (Baseline)
Week 20	Week 4
Week 24	Week 8
Week 28	Week 12
Week 32	Week 16
Week 36	Week 20
Week 40	Week 24
Week 44	Week 28
Week 48	Week 32
Week 52	Week 36

Appendix C. Partially Missing Date Imputation

Adverse Events

Partially missing AE start dates will be imputed in the ADaM dataset for AEs, but partially missing AE end dates will not be imputed in the same dataset. If the AE end date is complete with no missing year, month, or day, and the partially missing start date imputed by the rules below is after the AE end date, then the start date will be imputed by the AE end date.

Subjects who were treated with apremilast at Week 0/Visit 2

The principle of the imputation rules for subjects who were treated with apremilast initially is to treat the AE as treatment-emergent, i.e., occurring on or after the date of the first dose of IP, if possible.

Let an AE start date be represented as "$D_{Event}/M_{Event}/Y_{Event}$", and the date of the first dose of IP as "$D_{IP}/M_{IP}/Y_{IP}$". The following table gives the imputation rules for partially missing AE start dates for subjects who were treated with apremilast initially.

Table 3. Imputation Rules for Partially Missing AE Start Dates for Subjects Who Were Treated with Apremilast Initially

Scenario	Condition	Imputation Rule
Partially missing date includes year only (both month and day are missing)		
1	$Y_{Event} < Y_{IP}$	$12/31/Y_{Event}$
2	Otherwise, i.e., $Y_{IP} \leq Y_{Event}$	Max (date of first dose of IP, $1/1/Y_{Event}$)
Partially missing date includes both year and month (only day is missing)		
1	$Y_{Event} < Y_{IP}$, or ($Y_{Event} = Y_{IP}$ and $M_{Event} < M_{IP}$)	Last date of M_{Event}/Y_{Event}
2	Otherwise, i.e., $Y_{IP} < Y_{Event}$, or ($Y_{IP} = Y_{Event}$ and $M_{IP} \leq M_{Event}$)	Max (date of first dose of IP, $1/M_{Event}/Y_{Event}$)

Subjects who were treated with placebo at Week 0/Visit 2

The principle of the imputation rules for subjects who were treated with placebo initially and started apremilast treatment at Week 16 is to consider the AE starting on or after the date of the first dose of apremilast, if possible; if the partially missing start date suggests that it is prior to the date of the first dose of apremilast, the AE will be considered starting on or after the date of the first dose of IP, if possible.

The following are 4 scenarios considered in the imputation rules:

1. The partially missing AE start date suggests the date is prior to the date of the first dose of IP: impute it by the latest possible date (determined by the non-missing field of the date);
2. The partially missing AE start date suggests the date is after the date of the first dose of apremilast following Week 16: impute it by the earliest possible date (determined by the non-missing field of the date);
3. The partially missing AE start date is in the same year (if both month and day are missing), or the same year/month (if only day is missing) of the first dose of apremilast following Week 16: impute it by the date of the first dose of apremilast;
4. The partially missing AE start date suggests the date is no earlier than the date of the first dose of IP but prior to the date of the first dose of apremilast following Week 16: impute it by the date of the first dose of IP, or the earliest possible date (determined by the non-missing field of the date), whichever occurs later.

Let an AE start date be represented as "$D_{Event}/M_{Event}/Y_{Event}$", the date of the first dose of IP as "$D_{IP}/M_{IP}/Y_{IP}$", and the date of the first dose of apremilast following Week 16 as "$D_{APR}/M_{APR}/Y_{APR}$". The following table gives the imputation rules for partially missing AE start dates.

Table 4. Imputation Rules for Partially Missing AE Start Dates for Subjects Who Were Treated with Placebo Initially

Scenario	Condition	Imputation Rule
Partially missing date includes year only (both month and day are missing)		
1	$Y_{Event} < Y_{IP}$	$12/31/Y_{Event}$
2	$Y_{Event} > Y_{APR}$	$1/1/Y_{Event}$
3	$Y_{Event} = Y_{APR}$	Date of first dose of apremilast following Week 16
4	Otherwise, i.e., $Y_{IP} \leq Y_{Event} < Y_{APR}$	Max (date of first dose of IP, $1/1/Y_{Event}$)
Partially missing date includes both year and month (only day is missing)		
1	$Y_{Event} < Y_{IP}$, or ($Y_{Event} = Y_{IP}$ and $M_{Event} < M_{IP}$)	Last date of M_{Event}/Y_{Event}
2	$Y_{Event} > Y_{APR}$, or ($Y_{Event} = Y_{APR}$ and $M_{Event} > M_{APR}$)	$1/M_{Event}/Y_{Event}$

3	$Y_{Event} = Y_{APR}$ and $M_{Event} = M_{APR}$	Date of first dose of apremilast following Week 16
4	Otherwise, i.e., $Y_{IP} < Y_{Event} < Y_{APR}$, or ($Y_{IP} = Y_{Event} < Y_{APR}$ and $M_{IP} \leq M_{Event}$), or ($Y_{IP} = Y_{Event} = Y_{APR}$ and $M_{IP} \leq M_{Event} < M_{APR}$), or ($Y_{IP} < Y_{Event} = Y_{APR}$ and $M_{Event} < M_{APR}$)	Max (date of first dose of IP, $1/M_{Event}/Y_{Event}$)

Prior/Concomitant Medications/Procedures

Partially missing start/stop dates for prior/concomitant medications and partially missing start dates for prior/concomitant procedures will be imputed in the ADaM dataset for prior/concomitant medications/procedures. For prior/concomitant medications, if the stop date is complete with no missing year, month, or day, and the partially missing start date imputed by the rule below is after the stop date, then the start date will be imputed by the stop date.

Partially missing prior/concomitant medication/procedure start dates will be imputed by the earliest possible date given the non-missing field(s) of the date.

Partially missing prior/concomitant medication stop dates will be imputed by the latest possible date given the non-missing field(s) of the date.

Medical History

Partially missing medical history start dates will be imputed in the ADaM dataset for medical history (for the purposes of calculating durations of PsA and psoriasis). The 16th of the month will be used to impute a partially missing start date that has only the day missing, and July 1st will be used to impute a partially missing start date that has both the month and day missing.

Treatment Duration

Partially or completely missing last dose dates will be imputed in the ADaM dataset for treatment duration.

When partially missing last dose date is available, set last dose date to the maximum of [the earliest possible date given the non-missing field(s) of last dose date, the minimum of (the latest possible date given the non-missing field(s) of last dose date, last known date in database, first non-missing Early Termination (ET) visit date)]

When last dose date is completely missing, set last dose date to the minimum of (last known date in database, first non-missing Early Termination (ET) visit date)

Last known date in database is defined as maximum of (last visit date, lab, vital signs, ECG assessment date, AE start or end dates, concomitant medications start or end dates, concomitant procedure date, last dose date from 'Disposition- Treatment' page, treatment exposure start or end dates where doses were completely or partially taken, death date).

Appendix D. Reporting conventions

· Summary tabulation will be provided by treatment arm (e.g., Placebo vs. Apremilast for the Placebo-controlled Phase; Placebo/Apremilast vs. Apremilast /Apremilast for the Apremilast Extension Phase; Apremilast for Apremilast-exposure Period);

· Statistical test of the treatment difference will use two-sided significance level 0.05 for significance or nominal significance.

· P-values will be rounded to 4 decimal places. p-values that round to 0.0000 will be presented as '<0.0001' and p-values that round to 1.000 will be presented as '>0.9999';

· Confidence intervals (CIs) will be presented as two-sided 95% CIs unless specified differently in specific analysis;

· Summary statistics will consist of the number and percentage of subjects in each category for discrete variables, the sample size, mean, median, standard deviation (SD), minimum, and maximum for continuous variables, the 25th (Q1) and 75th (Q3) percentiles will also be applied to efficacy continuous variables;

· Mean and median values will be formatted to one more decimal place than the measured value. Standard deviation values will be formatted to two more decimal places than the measured value. Minimum and maximum values will be presented to the same number of decimal places as the measured value;

· Percentages will be rounded to one decimal place. The number and percentage of responses will be presented in the form xx (xx.x), where the percentage is in the parentheses;

· Analysis and summary tables will have the analysis population sample size (i.e., number of subjects) in the column header;

· Subject data listings will be provided to support the tables and graphs. Listings will be sorted for presentation in order of subject number and date of procedure or event.