

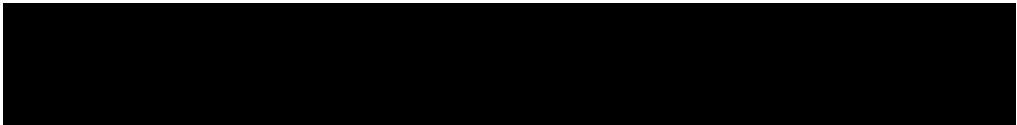
Statistical Analysis Plan

| | | |
|------------------------------|---|--|
| Protocol Title: | Phase 4, multicenter, randomized, double-blind, placebo-controlled, parallel-group study to evaluate the efficacy and safety of apremilast (CC-10004) in subjects with early, oligoarticular psoriatic arthritis despite initial stable treatment with either non-steroidal anti-inflammatory drugs (NSAIDs) and/or \leq 1 conventional synthetic disease-modifying antirheumatic drugs (csDMARD) | |
| Short Protocol Title: | Not Applicable | |
| Protocol Number: | 20200058 | |
| NCT Number: | NCT03747939 | |
| Authors: | [REDACTED] | |
| Sponsor: | AMGEN INC. One Amgen Center Drive, Thousand Oaks, CA 91320, United States | |
| SAP Date: | <u>Document Version</u> Original (v1.0) | <u>Date</u> 07 February 2023 |

| Version Number | Date (DDMMYYYY) | Summary of Changes, including rationale for changes |
|-----------------------|----------------------------|--|
| Original (v1.0) | 07FEB2023 | Not Applicable |

Table of Contents

| | |
|---|----|
| Table of Contents | 2 |
| 1. Introduction..... | 9 |
| 2. Objectives, Endpoints and Hypotheses..... | 9 |
| 2.1 Objectives and Endpoints/Estimands | 9 |
| 2.2 Hypotheses and/or Estimations..... | 17 |
| 3. Study Overview | 17 |
| 3.1 Study Design..... | 17 |
| 3.2 Sample Size | 19 |
| 3.3 Adaptive Design | 19 |
| 4. Covariates and Subgroups | 20 |
| 4.1 Planned Covariates..... | 20 |
| 4.2 Subgroups..... | 20 |
| 5. Definitions..... | 20 |
| 5.1 Derivations of Efficacy/ Health-related Quality of Life Endpoints | 20 |
| 5.1.1 Modified Minimal Disease Activity–Joints (MDA-Joints)..... | 20 |
| 5.1.2 Joint Assessment | 21 |
| 5.1.3 Patient’s Global Assessments of Disease Activity | 22 |
| 5.1.4 Patient’s Pain Visual Analogue Scale | 22 |
| 5.1.5 Data resource: Pain VAS is entered by the patient via tablets provided by ERT. HAQ-DI..... | 22 |
| 5.1.6 Enthesitis: Leeds Enthesitis Index | 23 |
| 5.1.7 BSA..... | 24 |
| 5.1.8 Physician’s Global Assessments of Disease Activity | 24 |
| 5.1.9 Enthesitis: [REDACTED] | 24 |
| 5.1.10 Dactylitis | 24 |
| 5.1.11 Clinical Disease Activity Index for Psoriatic Arthritis (cDAPSA)..... | 25 |
| 5.1.12 Psoriatic Arthritis Disease Activity Score (PASDAS)..... | 25 |
| 5.1.15 The PsA Impact of Disease 12-Item (PsAID-12) Questionnaire..... | 26 |

| | |
|--|---|
|  | |
| 5.2 | Derivations of Safety Endpoints30 |
| 5.2.1 | Treatment-emergent Adverse Events Leading to Drug Withdrawal, Leading to Drug Interruption, and Leading to Death, and Drug-related Treatment-emergent Adverse Events31 |
| 5.3 | Baseline Definitions.....31 |
| 5.4 | Time Points31 |
| 5.5 | Analysis Phases/Period.....32 |
| 6. | Analysis Sets34 |
| 6.1 | Full Analysis Set.....34 |
| 6.2 | Safety Analysis Set34 |
| 6.3 | Per Protocol Set(s).....34 |
| 6.4 | Health-related Quality-of-Life or Health Economics Analyses Set(s)34 |
| 6.5 | Pharmacokinetic/Pharmacodynamic Analyses Set(s).....35 |
| 6.6 | Interim Analyses Set(s)35 |
| 6.7 | Study-specific Analysis Sets.....35 |
| 6.7.1 | Apremilast Subjects as Randomized or Transitioned Population35 |
| 6.7.2 | Apremilast Subjects as Treated Population35 |
| 7. | Planned Analyses35 |
| 7.1 | Interim Analysis and Early Stopping Guidelines35 |
| 7.2 | Primary Analysis35 |
| 7.3 | Final Analysis.....36 |
| 8. | Data Screening and Acceptance.....36 |
| 8.1 | General Principles.....36 |
| 8.2 | Data Handling and Electronic Transfer of Data36 |
| 8.3 | Handling of Missing and Incomplete Data36 |
| 8.3.1 | Adverse Events36 |
| 8.3.2 | Prior/Concomitant Medications/Procedures38 |
| 8.3.3 | Medical History.....39 |
| 8.4 | Detection of Bias39 |
| 8.5 | Outliers40 |
| 8.6 | Distributional Characteristics40 |
| 8.7 | Validation of Statistical Analyses.....40 |
| 9. | Statistical Methods of Analysis.....40 |
| 9.1 | General Considerations.....40 |
| 9.1.1 | Reporting and Analysis Conventions40 |
| 9.1.2 | Statistical Methodology for Efficacy41 |

| | | |
|--------|--|----|
| 9.1.3 | Multiplicity..... | 42 |
| 9.2 | Subject Accountability | 42 |
| 9.3 | Important Protocol Deviations | 43 |
| 9.4 | Demographic and Baseline Characteristics | 43 |
| 9.4.1 | Demographics | 44 |
| 9.4.2 | Baseline or Disease Characteristics | 44 |
| 9.4.3 | Medical History | 47 |
| 9.4.4 | Prior and Concomitant Procedures..... | 47 |
| 9.4.5 | Prior and Concomitant Medications | 47 |
| 9.5 | Efficacy Analyses | 49 |
| 9.5.1 | Analyses of Primary Efficacy Endpoint | 49 |
| 9.5.2 | Analyses of Secondary Efficacy Endpoint(s)..... | 51 |
| 9.5.3 | Analyses of Exploratory Efficacy Endpoint(s)..... | 52 |
| 9.6 | Safety Analyses | 54 |
| 9.6.1 | Analyses of Primary Safety Endpoint..... | 54 |
| 9.6.2 | Adverse Events | 55 |
| 9.6.3 | Adverse Events of Special Interest | 57 |
| 9.6.4 | Laboratory Test Results | 57 |
| 9.6.5 | Vital Signs | 59 |
| 9.6.6 | Physical Measurements | 59 |
| 9.6.7 | Electrocardiogram | 59 |
| 9.6.8 | Antibody Formation | 59 |
| 9.6.9 | Exposure to Investigational Product | 59 |
| 9.6.10 | Exposure to Non-investigational Product..... | 61 |
| 9.6.11 | Exposure to Other Protocol-required Therapy | 61 |
| 9.7 | Other Analyses | 61 |
| 9.7.1 | Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints..... | 61 |
| 9.7.2 | Analyses of Clinical Outcome Assessments | 61 |
| 9.7.3 | Analyses of Health Economic Endpoints | 61 |
| 9.7.4 | Analyses of Biomarker Endpoints | 61 |
| 10. | Changes From Protocol-specified Analyses..... | 62 |
| 11. | Literature Citations / References..... | 62 |
| 12. | Prioritization of Analyses..... | 62 |
| 13. | Data Not Covered by This Plan..... | 62 |
| 14. | Appendices..... | 63 |
| | Appendix A. Analytical Windows..... | 64 |
| | Appendix B. Marked Abnormality Criteria | 66 |

List of Tables

| | |
|--|----|
| Table 1. Response criteria for the PASDAS | 26 |
| Table 2. Imputation Rules for Partially Missing AE Start Dates for Subjects Who Were Treated with Apremilast Initially..... | 37 |
| Table 3. Imputation Rules for Partially Missing AE Start Dates for Subjects Who Were Treated with Placebo Initially..... | 38 |
| Table 4. Normal Ranges for Shift Table | 59 |
| Table 5. Table for Visit Mapping for by Time Point Analysis | 64 |
| Table 6. Marked Abnormalities Criteria | 66 |

List of Figures

No table of figures entries found.

List of Abbreviations

| Abbreviation | Explanation |
|--------------|--|
| AE | Adverse Event |
| ATC | Anatomical Therapeutic Chemical |
| ATC | Anatomical-Therapeutic-Chemical |
| BID | Twice Daily |
| | |
| BSA | Body Surface Area (%) Affected by Psoriasis |
| cDAPSA | Clinical Disease Activity in Psoriatic Arthritis |
| CI | Confidence Interval |
| CMH | Cochran-Mantel-Haenszel |
| COVID-19 | Coronavirus Disease 2019 |
| CRF or eCRF | (electronic) Case Report Form |
| CRO | Contract Research Organization |
| CRP | C-Reactive Protein |
| csDMARD | Conventional Synthetic Disease-Modifying Antirheumatic Drugs |
| CSR | Clinical Study Report |
| EAIR | Exposure-Adjusted Incidence Rate |
| EE | Early Escape |
| | |
| ERT | eResearchTechnology, Inc |
| GSO-DM | Global Study Operations-Data Management |
| GSP | Global Statistical Programming |
| HAQ-DI | Health Assessment Questionnaire – Disability Index |
| | |
| hsCRP | High Sensitivity C-Reactive Protein |
| IA | Intra-Articular |
| IP | Investigational Product |
| IPDs | Important Protocol Deviations |
| IWRS | Interactive Web Response System |
| L/R | Left/Right |
| | |
| LEI | Leeds Enthesitis Index |
| LN | Natural Logarithm |

| | |
|----------|---|
| LS Mean | Least Squares Mean |
| MAR | Missing At Random |
| MCID | Minimal Clinically Important Differences |
| MDA | Minimal Disease Activity |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MMRM | Mixed-effect Model for Repeated Measures |
| MNAR | Missing Not At Random |
| MTX | Methotrexate |
| NR | Non-Responder |
| NRI | Non-responder Imputation |
| NRS | Numeric Rating Scale |
| NSAIDs | Non-Steroidal Anti-Inflammatory Drugs |
| PASDAS | Proportion Of Patients Achieving A Good Or Moderate Response |
| | |
| PD | Pharmacodynamic |
| PP | Per-Protocol (Set) |
| PROs | Patient-Reported Outcomes |
| PsA | Psoriatic Arthritis |
| PsAID-12 | Psoriatic Arthritis Impact Of Disease 12-Item For Clinical Trials |
| PT | Preferred Term |
| QOL | Quality Of Life |
| SAP | Statistical Analysis Plan |
| SD | Standard Deviation |
| | |
| SI | Standard International (unit) |
| SJC | Swollen Joint Count |
| SOC | System Organ Class |
| sPGA | static Physician Global Assessment |
| | |
| SSZ | Sulfasalazine |
| TEAE | Treatment-emergent Adverse Event |
| TJC | Tender Joint Count |
| VAS | Visual Analogue Scale |

| | |
|-------|---|
| VLDA | Very Low Disease Activity |
| 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 CC-10004-PSA-013 amendment 5 for study 20200058, AMG407 dated 12 July 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.

The SAP will be finalized and signed prior to the unblinding of the Week 24 database which will be the primary analysis database. All statistical analyses detailed in this SAP will be conducted using SAS® Version 9.4 or higher.

2. Objectives, Endpoints and Hypotheses

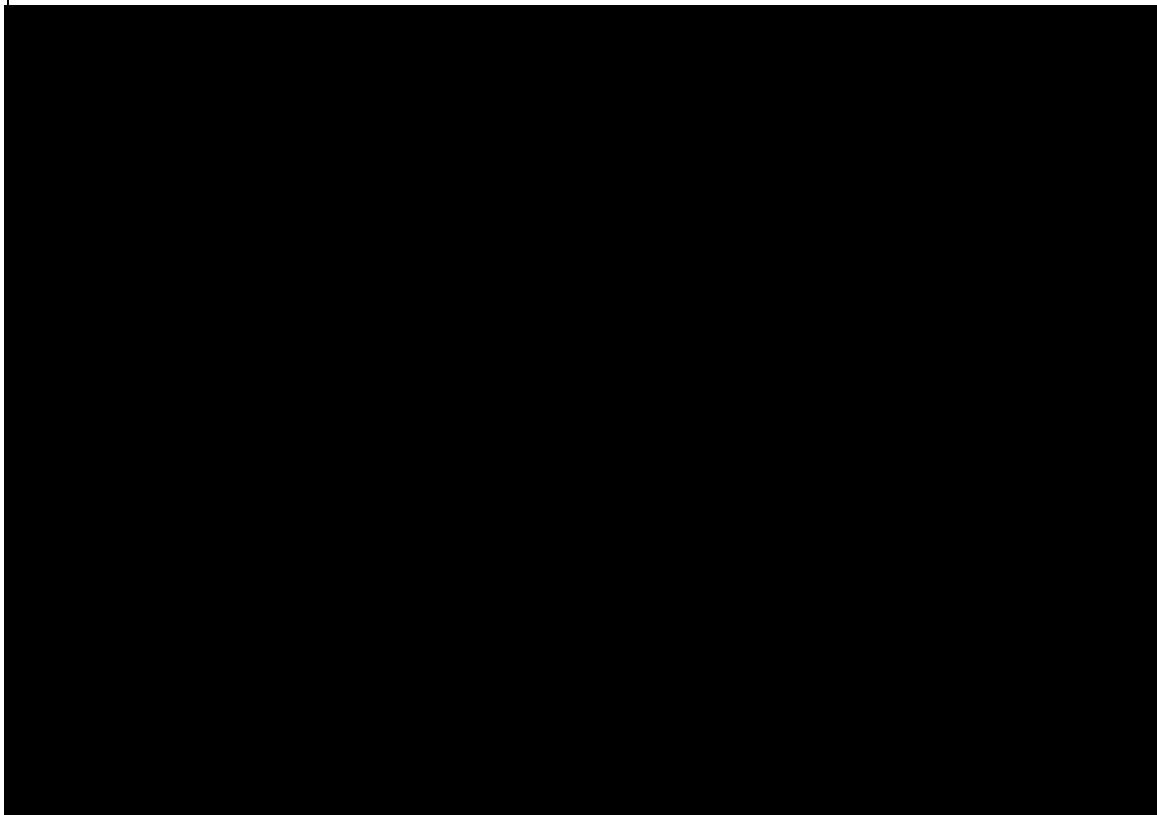
2.1 Objectives and Endpoints/Estimands

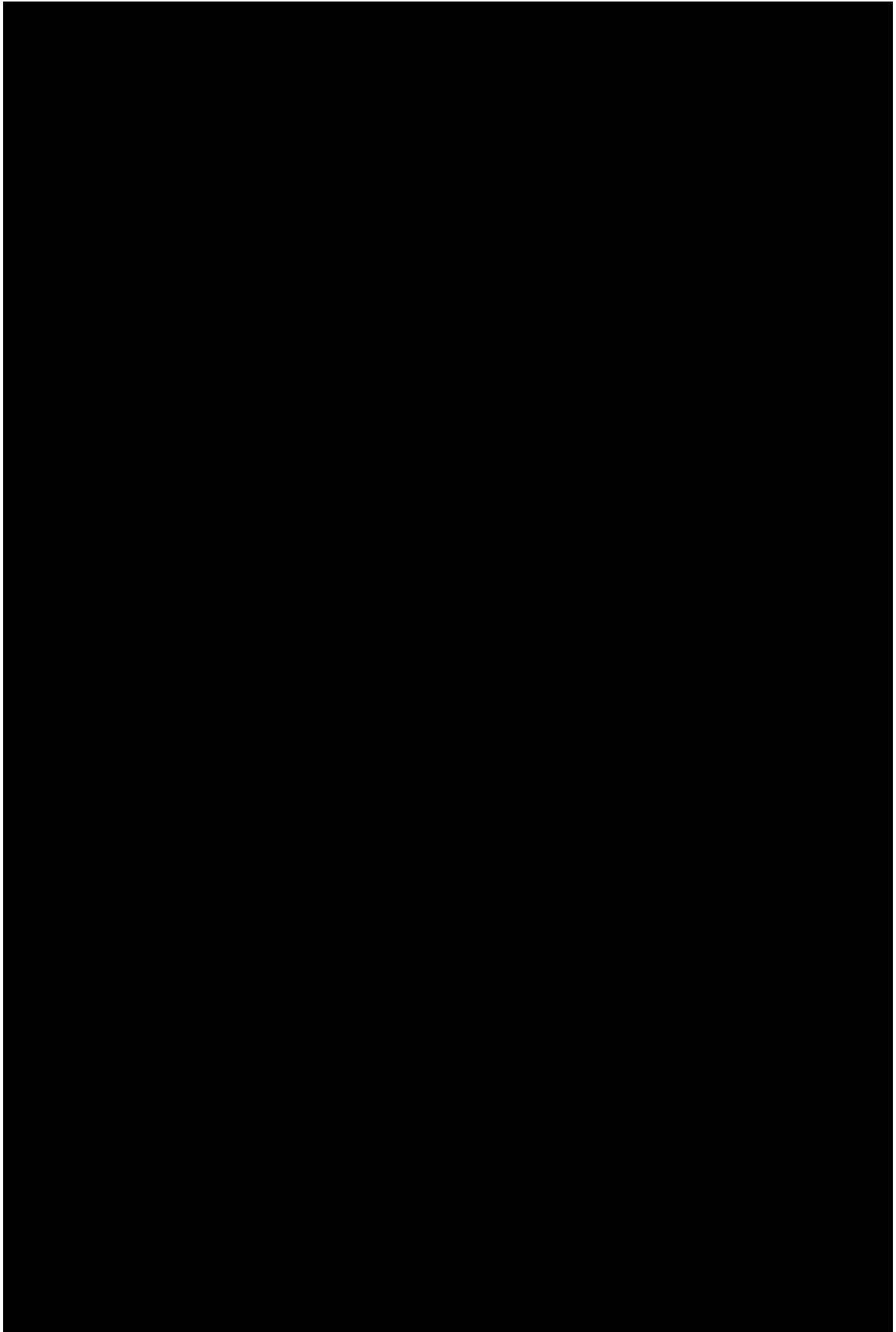
| Objectives | Endpoints |
|--|--|
| Primary | |
| <ul style="list-style-type: none"> To evaluate the efficacy of apremilast 30 mg BID \pm NSAIDs and/or csDMARDs vs. Placebo \pm NSAIDs and/or csDMARDs in subjects with early oligoarticular psoriatic arthritis (PsA), assessed by modified MDA (MDA-Joints). | <ul style="list-style-type: none"> Proportion of subjects who achieve a clinical state of minimal disease activity (MDA-Joints) at Week 16, which is defined as achieving $SJC \leq 1$ and $TJC \leq 1$ (based on sentinel joints), plus 3 out of the following 5 criteria: <ul style="list-style-type: none"> Psoriasis BSA $\leq 3\%$; Patient pain (VAS) ≤ 15; Patient global disease activity (VAS) ≤ 20; HAQ-DI ≤ 0.5; Tender enthesal points ≤ 1 (based on Leeds Enthesitis Index - LEI) |
| Secondary | |
| <ul style="list-style-type: none"> To evaluate the impact of treatment with apremilast 30 mg BID \pm NSAIDs and/or csDMARDs vs. Placebo \pm NSAIDs and/or csDMARDs on disease activity in | <ul style="list-style-type: none"> cDAPSA remission or low disease activity: Proportion of subjects who achieve remission (defined by clinical disease activity in psoriatic arthritis [cDAPSA] ≤ 4 score) or low disease activity |

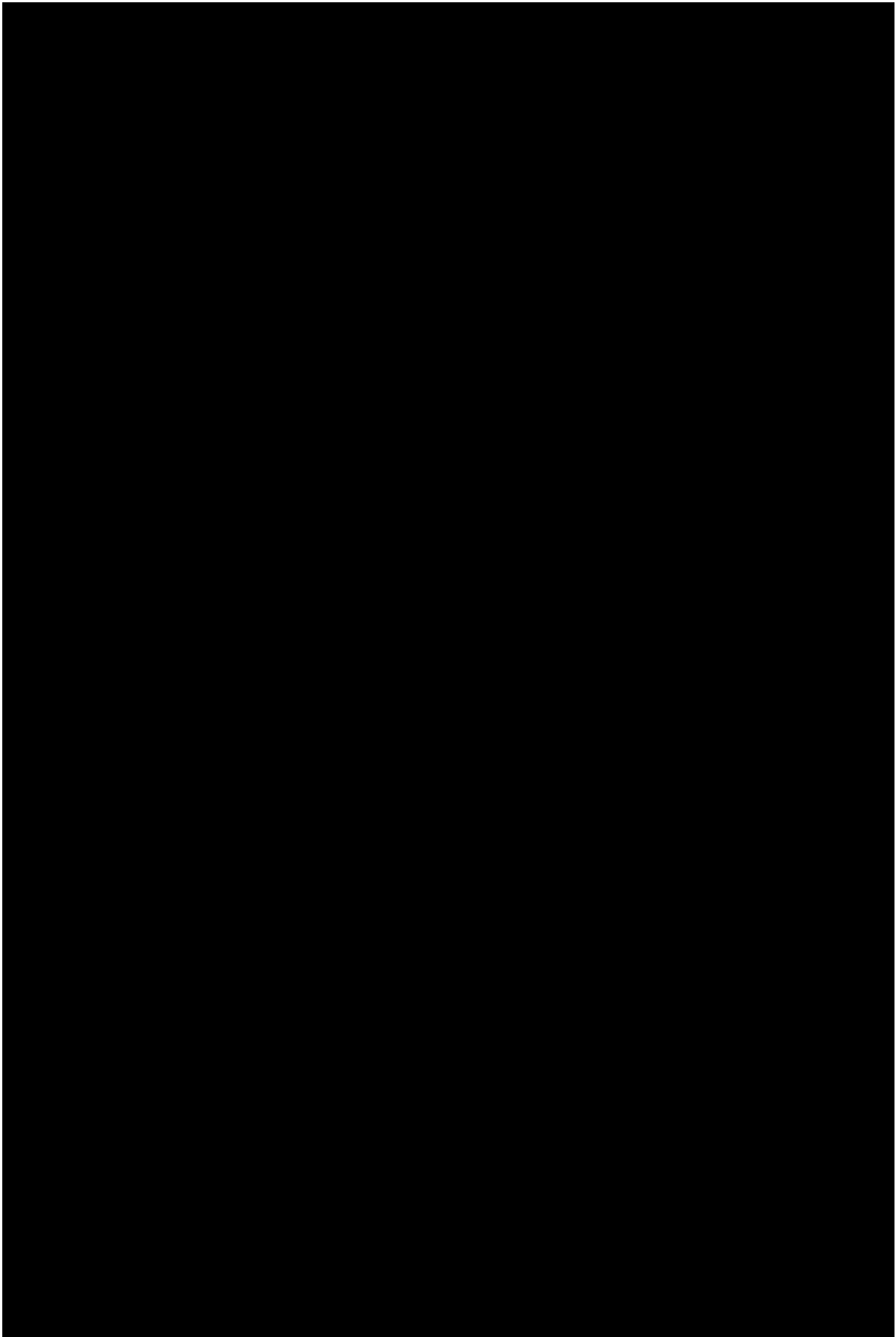
| | |
|---|--|
| <p>subjects with early oligoarticular PsA.</p> | <p>(defined by cDAPSA > 4 but ≤ 13 score) at week 16.</p> <ul style="list-style-type: none"> • Swollen Joint Count (SJC) ≤ 1: Proportion of subjects with SJC ≤ 1 at week 16, based on sentinel joints. • Tender Joint Count (TJC) ≤ 1 : Proportion of subjects with TJC ≤ 1 at week 16, based on sentinel joints. • Patient's global assessment of disease activity: Proportion of subjects whose global assessment of disease activity score ≤ 20 in the visual analogue scale (VAS) at week 16. • Patient's assessment of pain: Proportion of subjects whose assessment of pain score ≤ 15 in VAS at week 16. • PASDAS good and moderate response: Proportion of patients achieving a good or moderate response in PASDAS score (Table 1) at week 16. |
| <ul style="list-style-type: none"> • To evaluate the impact of apremilast 30 mg BID ± NSAIDs and/or csDMARDs vs. Placebo ± NSAIDs and/or csDMARDs on patient-reported outcomes (PROs) in subjects with early oligoarticular PsA. | <ul style="list-style-type: none"> • PsAID-12: Change from Baseline in the Psoriatic Arthritis Impact of Disease 12-item for clinical trials (PsAID-12) questionnaire at week 16. |

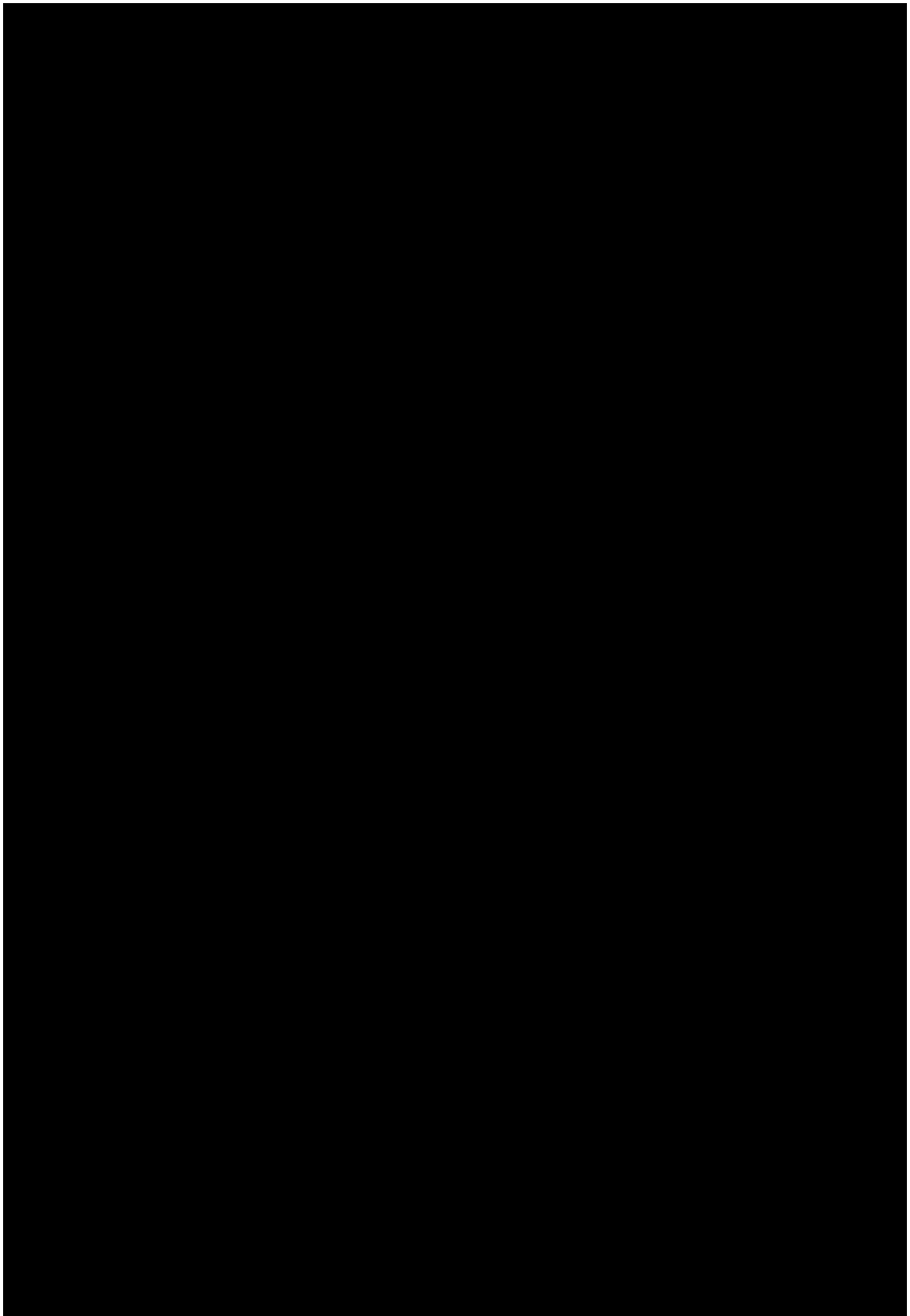
| | |
|---|--|
| <ul style="list-style-type: none">To evaluate the safety and tolerability of apremilast 30 mg BID \pm NSAIDs and/or csDMARDs vs. Placebo \pm NSAIDs and/or csDMARDs in subjects with early oligoarticular PsA | <ul style="list-style-type: none">Adverse events: Type, frequency, severity, and relationship of adverse events (AEs) to investigational product (IP).Discontinuation due to AEs: Number of subjects who discontinue IP due to any AE.Clinically significant changes in physical examination, vital signs, and/or laboratory findings: Frequency of clinically significant changes in physical examination, vital signs, and/or laboratory findings. |
|---|--|

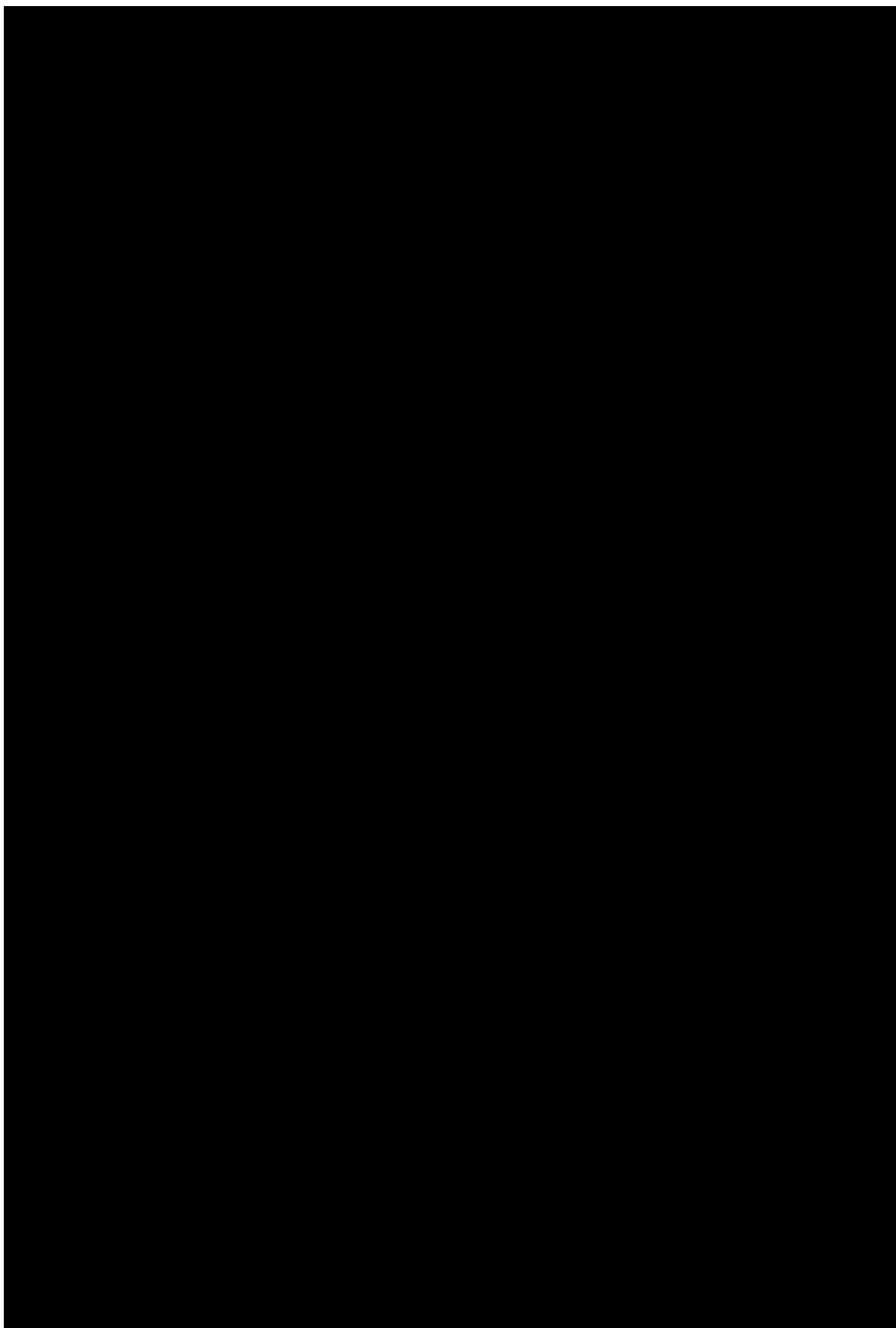
Exploratory

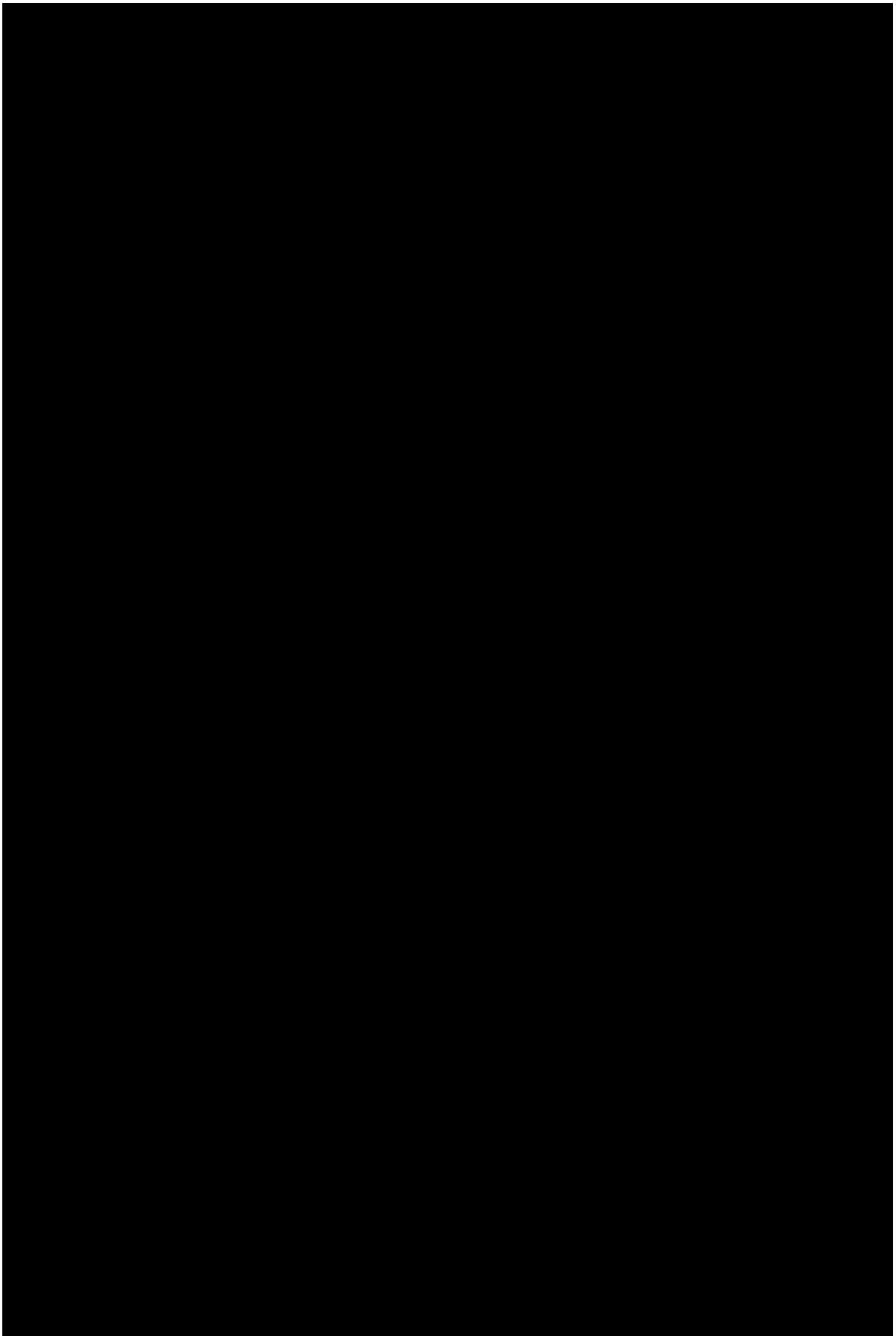


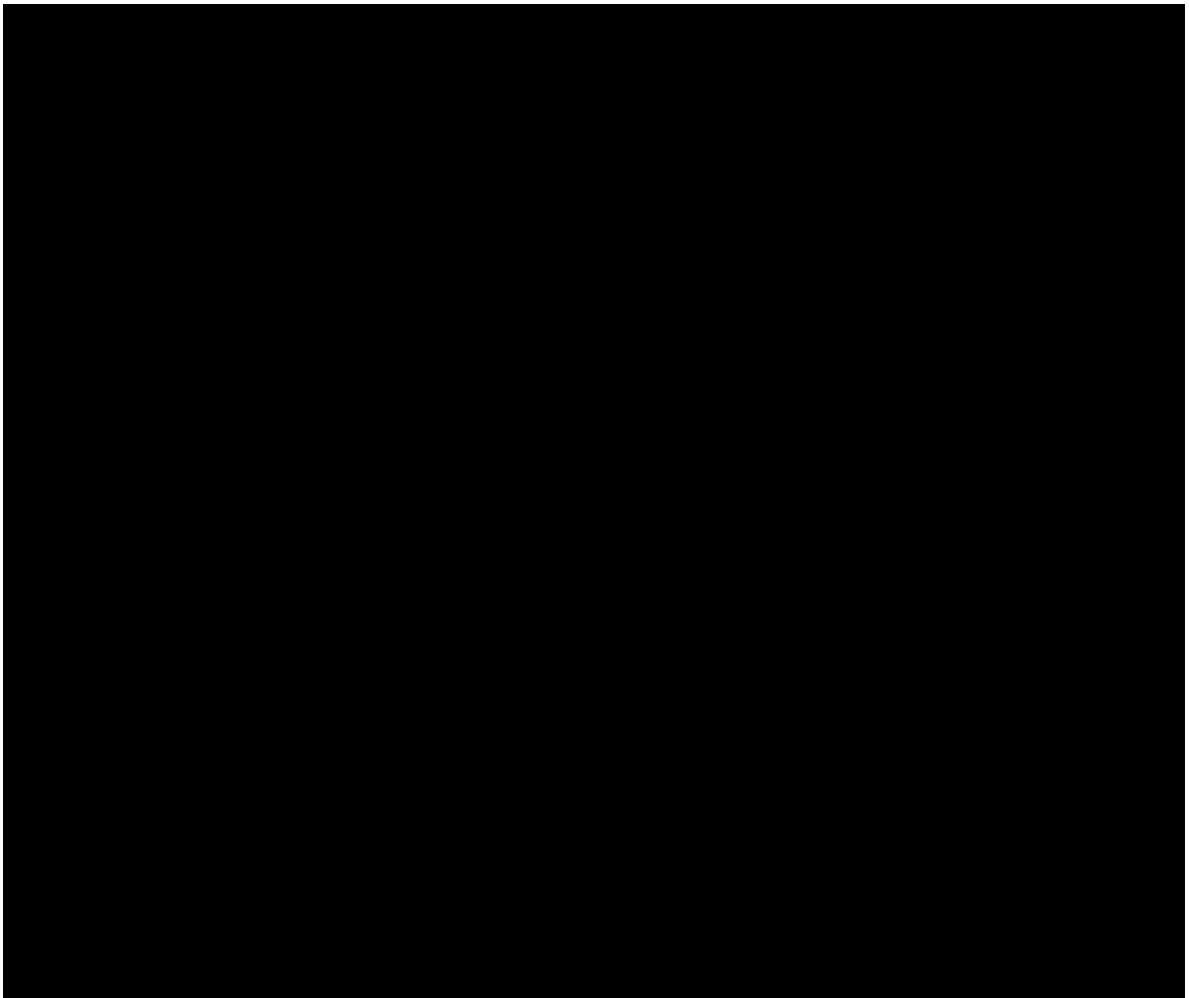












2.2 Hypotheses and/or Estimations

The primary hypothesis for this study is: Apremilast \pm NSAIDs and/or csDMARDs is superior to placebo \pm NSAIDs and/or csDMARDs in subjects with early oligoarticular psoriatic arthritis as measured by proportion of subjects achieving MDA-Joints.

H_0 : the difference in proportion of subjects who achieved a clinical state of minimal disease activity (MDA-Joints) at Week 16 in apremilast and in that of placebo = 0.

H_1 : the difference in proportion of subjects who achieved a clinical state of minimal disease activity (MDA-Joints) at Week 16 in apremilast and in that of placebo \neq 0.

3. Study Overview

3.1 Study Design

This Phase 4 multicenter, randomized, placebo-controlled, double-blind study is designed to evaluate the efficacy, safety and tolerability of apremilast over 48 weeks of treatment in approximately 285 subjects with early oligoarticular psoriatic arthritis.

subjects will be randomized in a 2:1 ratio to either apremilast 30 mg BID or identically appearing placebo via an Interactive Web Response System (IWRS).

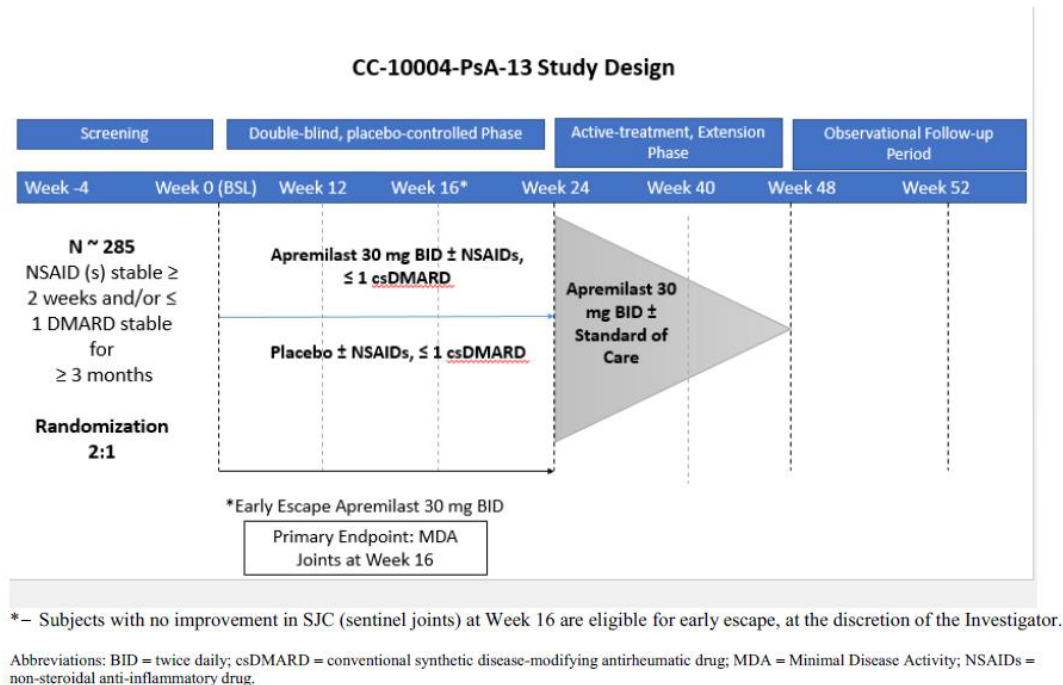
Randomization to apremilast arm or placebo arm will be stratified by concomitant use of glucocorticosteroids at baseline (yes/no) and prior/concomitant use of a csDMARD, either MTX or SSZ, (i.e., csDMARD-naïve; csDMARD use prior to baseline [treatment not continued]; or csDMARD use prior to and concomitant at baseline). The number of subjects who are csDMARD-naïve is targeted to comprise up to 50 % of the subjects enrolled in the study. There is no planned protocol extension following the end of the study.

This is a 56-week study which will consist of 4 phases:

- Screening Phase – up to 4 weeks
- (Randomized, Double-blind) Placebo-controlled (Treatment) Phase – Weeks 0 to 24.
 - Stratified randomization (2:1 ratio to either apremilast 30 mg BID or placebo).
 - Dose-titrated.
 - Subjects with no improvement in SJC (based on sentinel joints, i.e., the joints affected at baseline) at Week 16 are eligible for Early Escape (EE), at the discretion of the Investigator; Placebo-treated subjects who escape early will be transitioned to apremilast 30 mg BID in a blinded fashion via IWRS, while apremilast-treated subjects will remain on the original treatment assigned at baseline.
 - The patients remaining on placebo will be transitioned to apremilast 30 mg BID at Week 24.
 - All subjects who complete this 24-week treatment phase will have the option to enter the Active-treatment Extension Phase.
- (Active-treatment) Extension Phase – Week 24 to Week 48.
 - Beginning at Week 24 Visit, all subjects in the Extension Phase will be dispensed open-label apremilast (30 mg BID).
 - All subjects will receive apremilast 30 mg BID until the completion of the study (i.e., up to Week 48 Visit)
- (Post-treatment) Observational Follow-up Phase – up to 4 weeks.

- All subjects who complete the Placebo-controlled Phase or the Extension Phase or discontinue early will participate in the 4-week Observational Follow-up Phase.
- The Observational Follow-up visit should be performed 4 weeks after the last dose of IP.

The study schematic is presented in the following figure:



3.2 Sample Size

Sample size estimation was based on the results from integrated analyses of subpopulation data from five Phase 3 studies (PSA-002, PSA-003, PSA-004, PSA-005, and PSA-006). Assuming a 15% dropout rate by the primary endpoint, a sample size of [REDACTED] for the Apremilast 30 mg BID and placebo groups, respectively, will have 80% power to detect a true 15% difference (30% versus 15%) in the proportions of subjects achieving the modified MDA (MDA-Joints) between the Apremilast 30 mg BID and placebo groups, using a chi-square test with a two-sided significance level of 0.05. The calculation was implemented using nQuery 7.0.

3.3 Adaptive Design

Not Applicable.

4. Covariates and Subgroups

4.1 Planned Covariates

Covariates includes but not be limited to baseline value for each measured endpoint.

4.2 Subgroups

To determine whether the treatment effect is consistent across various subgroups, the primary analysis of the primary endpoint will be repeated for each category, as appropriate, of the following classification variables:

Subgroups

- Sex (male, female)
- Race (White, non-White)
- Age (< 65, ≥ 65 years)
- Baseline weight (< 70, ≥ 70 to < 85, ≥ 85 to < 100, ≥ 100 kg)

- Region (North America, Europe, Rest of the World [including Russia])
- Baseline glucocorticosteroid use (yes, no)
- Prior/concomitant use of csDMARD, either MTX or SSZ, (naïve, prior use only, prior and concomitant use)
- Prior use of csDMARD (yes, no)
- Concomitant use of csDMARD (yes, no).

The subgroup categories proposed above for the individual demographic and baseline parameters may be adjusted as appropriate. Additionally, the MDA-Joints response rates at Week 16 will be described graphically with forest plots by subgroup.

5. Definitions

5.1 Derivations of Efficacy/ Health-related Quality of Life Endpoints

Baseline definition for all efficacy endpoints is given in Section 5.3. Change from baseline is calculated as on-treatment value minus the baseline value. Handling of time points is described in Section 5.4.

5.1.1 Modified Minimal Disease Activity–Joints (MDA-Joints)

For this study, the MDA-Joints (Appendix C of Protocol) is the primary endpoint and is defined as Tender Joint Counts (TJC) ≤ 1 and Swollen Joint Counts (SJC) ≤ 1 plus 3 of

the following 5 criteria: 1) psoriasis Body Surface Area (BSA) $\leq 3\%$, 2) patient pain VAS on a 100-mm scale ≤ 15 , 3) patient's global assessment on a 100-mm scale ≤ 20 , 4) physical function [HAQ-DI] ≤ 0.5 , and 5) enthesitis count ≤ 1 based on the Leeds Enthesitis Index. Joint assessment will be based on 66 swollen joints and 68 tender joints. Assessment of tenderness and swelling of the individual joints will be performed by the evaluator; however, calculation of the MDA-Joints will be done programmatically.

Data resource: each component is entered by the physician or the subject via tablets provided by eResearchTechnology, Inc (ERT). The binary value of meeting MDA-Joints criteria or not is calculated in the analysis dataset.

5.1.2 Joint Assessment

TJC or SJC will be derived from the joint assessments entered by the physician via tablets provided by ERT. Joint assessment for tenderness and swelling will be performed for 68 and 66 joints (Appendix D of Protocol), respectively, at every visit. Joint tenderness and swelling will be noted in the following categories, with no quantitation of severity, as shown on the tablet.

- Pain/Tenderness Only
- Swelling Only
- Pain/Tenderness and Swelling
- Permanently Not Assessable
- Temporarily Not Assessable Due to IA Injection
- Temporarily Not Assessable - Other

In the event that a joint is injected with intra-articular (IA) glucocorticoid (allowed only after week 24 assessments are completed, and up to the Week 48 Visit), this joint should be marked as "Temporarily Not Assessable Due to IA Injection" if the study visit occurs within 4 weeks (28 days) of the IA injection. In order to maintain consistency throughout the study, the same evaluator should perform the joint assessments for a given subject at a study site at each study visit.

As a joint assessment may be recorded as "temporarily not assessable" or "permanently not assessable", the following data handling rules will be applied to handle the data in the calculation of SJC and TJC.

- Joints temporarily or permanently not assessable at baseline are excluded from joint count throughout the study.

- For other un-assessed joints at baseline, the joint assessment at the screening visit, if assessed, is used as the baseline assessment; otherwise, the joint is excluded from joint count throughout the study.
- For joints that become either temporarily or permanently not assessable at a post-baseline visit, the last observed joint assessment (including baseline and post-baseline assessments) will be used for that visit.

After applying these data handling rules, TJC will be calculated by adding the number of joints with outcomes of 'Pain/Tenderness Only' or 'Pain/Tenderness and Swelling'. SJC will be calculated by adding the number of joints with outcomes of 'Swelling Only' or 'Pain/Tenderness and Swelling'. TJC and SJC will be derived based on the sentinel joints and all joints (66 joints for swelling and 68 joints for tenderness).

5.1.3 Patient's Global Assessments of Disease Activity

The Patient's (Subject's) Global Assessments of Disease Activity (Appendix E of Protocol) is an assessment of how active a subject's psoriatic arthritis was on average during the last week. The subject will be asked to place a vertical line on a visual analogue scale (VAS) on which the left-hand boundary represents the lowest level of disease activity and the right-hand boundary represents the highest. Distance from the mark to the left-hand boundary will be recorded in millimeters, except when calculating cDAPSA score, where the distance will be calculated in centimeters.

Data resource: This VAS is entered by the patient via tablets provided by ERT.

5.1.4 Patient's Pain Visual Analogue Scale

The pain in visual analogue scale (VAS) (Appendix E of Protocol) is the subject's assessment of how much pain he/she had, on average, last week in his/her joints due to PsA. The subject will be asked to place a vertical line on a VAS on which the left-hand boundary represents "no pain" and the right-hand boundary represents "pain as severe as can be imagined". The distance from the mark to the left-hand boundary will be recorded in millimeters, except when calculating cDAPSA score, where the distance will be calculated in centimeters.

5.1.5 Data resource: Pain VAS is entered by the patient via tablets provided by ERT. HAQ-DI

The HAQ-DI (Appendix F of Protocol) is a 20-question, self-administered instrument that measures the subject's functional ability on a 4-level difficulty scale (0 to 3, with 0

representing normal or no difficulty; and 3 representing an inability to perform). Eight categories of functioning are included: dressing, rising, eating, walking, hygiene, reach, grip, and usual activities. This scale is sensitive to change and is a good predictor of future disability. The HAQ-DI score will be set to missing if a subject does not have values for at least 6 of the 8 categories.

Computing the HAQ-DI Score will follow the 2 steps below:

1. Find the highest score for each category. If use of aids/devices and/or help from another person are indicated, adjust the highest score for the associated category by increasing a lower score, which is between 0 and 2, to represent underlying disability more accurately. For example, if use of aids/devices and/or help from another person are indicated, adjust the highest score to 2 if the highest score is a 0 or a 1. If the highest score is 2 or 3, retain the value.
2. Sum the highest scores from all categories answered and divide it by the number of categories answered (must be a minimum of 6) to obtain an HAQ-DI score of 0-3 (3=worst functioning).

Please note, only one score will be calculated per HAQ-DI assessment.

If the “other” category is checked on the HAQ-DI assessment, then this is expected to be an aide not associated for any of the existing categories, thus this will be excluded from the HAQ-DI score calculation.

Data resource: HAQ-DI is entered by the patient via tablets provided by ERT.

5.1.6 Enthesitis: Leeds Enthesitis Index

In the evaluation of MDA-Joints, enthesitis will be assessed by the Leeds Enthesitis Index (LEI) (Appendix G of Protocol).

For the LEI, pain will be assessed at the following 6 entheses (tendon insertions) sites: Lateral epicondyle humerus (left/right (L/R)); Medial condyle femur (L/R); and Achilles tendon insertion into calcaneum (L/R). Tenderness at each site is quantified on a dichotomous basis: 0 means nontender and 1 means tender. The LEI ranges from 0 to 6.

Data resource: LEI is entered by the physician via tablets provided by ERT.

5.1.7 BSA

BSA (Appendix H of Protocol) is a measurement of involved skin. The overall BSA affected by psoriasis is estimated based on the palm area of the subject's hand (entire palmar surface or "handprint"), which equates to approximately 1% of total body surface area. All subjects will have their BSA assessed by the Evaluator.

Data resource: BSA is entered by the physician via tablets provided by ERT.

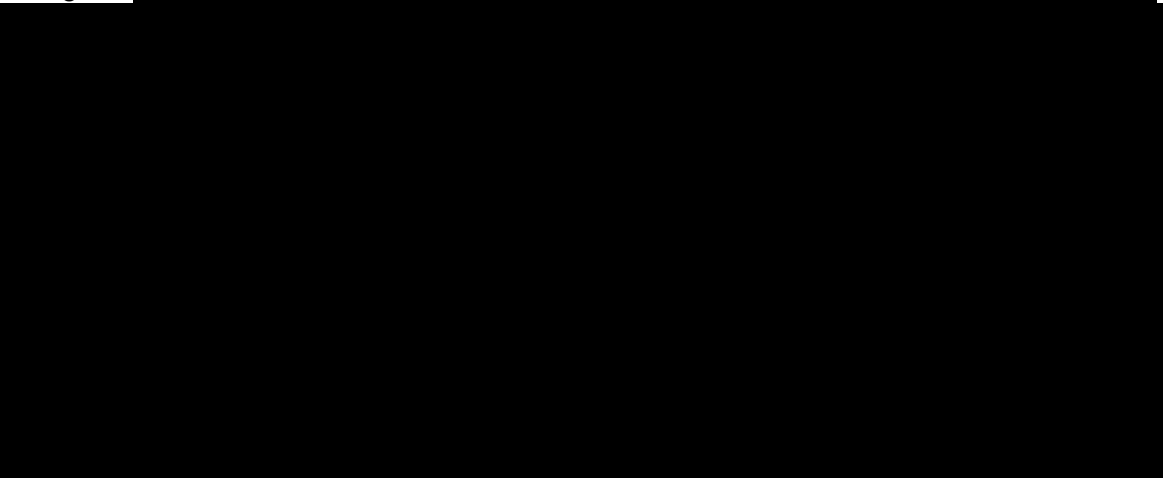
5.1.8 Physician's Global Assessments of Disease Activity

The Physician's (Evaluator's) Global Assessments of Disease Activity is an assessment of how active a subject's psoriatic arthritis was on average during the last week. For the Physician's (Evaluator's) Global Assessments of Disease Activity, evaluator should mark a single vertical mark on a 100-mm VAS, with zero representing the lowest level of disease activity or pain and 100 representing the highest. The distance from the mark to the left-hand boundary of the scale will be recorded in millimeters.

Data resource: This VAS is entered by the physician via tablets provided by ERT.

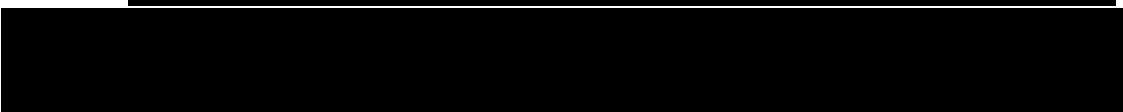
5.1.9 Enthesitis

In addition to assessment by the Leeds Enthesitis Index, enthesitis will also be assessed using the



5.1.10 Dactylitis

Dactylitis assessment will be done at screening, only a binary evaluation is needed (yes/no).



5.1.11 Clinical Disease Activity Index for Psoriatic Arthritis (cDAPSA)

The cDAPSA score (Appendix K of Protocol) is based on the numerical summation of 4 disease activity variables: tender and swollen joints, subject's (patient's) global assessments of disease activity (SGA) and Subject's assessment of pain on a 10 cm Visual Analogue Scale (VAS).

cDAPSA = SJC(66)+TJC(68)+SGA(VAS 0-10 cm)+Pain(VAS 0-10 cm), with range of 0-154, and it will be set to missing if any of the 4 components is missing. A higher score indicates greater disease activity. Disease activity based on cDAPSA will be categorized as remission if cDAPSA ≤ 4 , low disease activity if $4 < \text{cDAPSA} \leq 13$, moderate disease activity if $13 < \text{cDAPSA} \leq 27$, and high disease activity if cDAPSA > 27 . The calculation of the cDAPSA will be performed programmatically.

5.1.12 Psoriatic Arthritis Disease Activity Score (PASDAS)

The PASDAS (Appendix L of Protocol) is a weighted index comprising assessments of joints, function, acute-phase response, quality of life (QOL), and patient and physician by VAS. The calculation of the PASDAS will be performed programmatically using the formula below:

$$\text{PASDAS} = ((0.18 \times \sqrt{\text{Physician global VAS}}) + (0.159 \times \sqrt{\text{Patient global VAS}}) - (0.253 \times \sqrt{\text{SF36-PCS}}) + (0.101 \times \text{LN}(\text{Swollen joint count} + 1)) + (0.048 \times \text{LN}(\text{Tender joint count} + 1)) + (0.23 \times \text{LN}(\text{Leeds Enthesitis Count} + 1)) + (0.377 \times \text{LN}(\text{Dactylitis count} + 1)) + (0.102 \times \text{LN}(\text{CRP} + 1) + 2)) \times 1.5.$$

Note: LN = natural logarithm, SF36 = Medical Outcomes Study Short Form-36, SF36-PCS = physical component summary scale of SF36, CRP = C-reactive protein in mg/L. High sensitivity C-reactive protein (hsCRP) and rheumatoid factor will be analyzed at a central laboratory.

All VAS scores are 0–100 mm. Swollen joint count is based on 66 joints, and tender joint count based on 68 joints. The score range of the PASDAS is 0–10, with worse disease

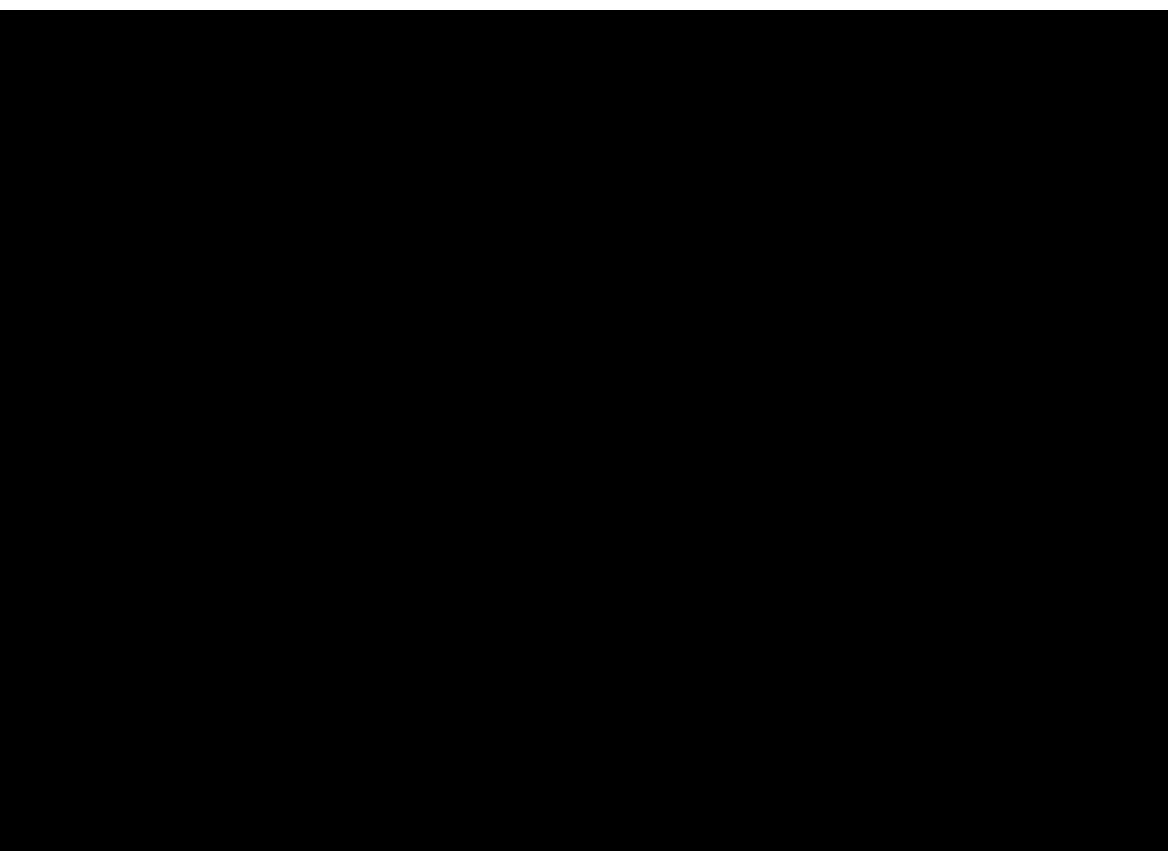
activity represented by higher scores. If CRP is reported as below detection limit, then the value of detection limit will be used for PASDAS calculation.

If one or more components are missing, then PASDAS is set to be missing.

Table 1. Response criteria for the PASDAS

| Final PASDAS Score | Improvement | | |
|---------------------|-------------|---------------------|------------|
| | ≥ 1.6 | < 1.6 but > 0.8 | ≤ 0.8 |
| ≤ 3.2 | 1 | 2 | 3 |
| > 3.2 but < 5.4 | 2 | 2 | 3 |
| ≥ 5.4 | 2 | 3 | 3 |

1= good response; 2= moderate response; 3=poor response



5.1.15 The PsA Impact of Disease 12-Item (PsAID-12) Questionnaire

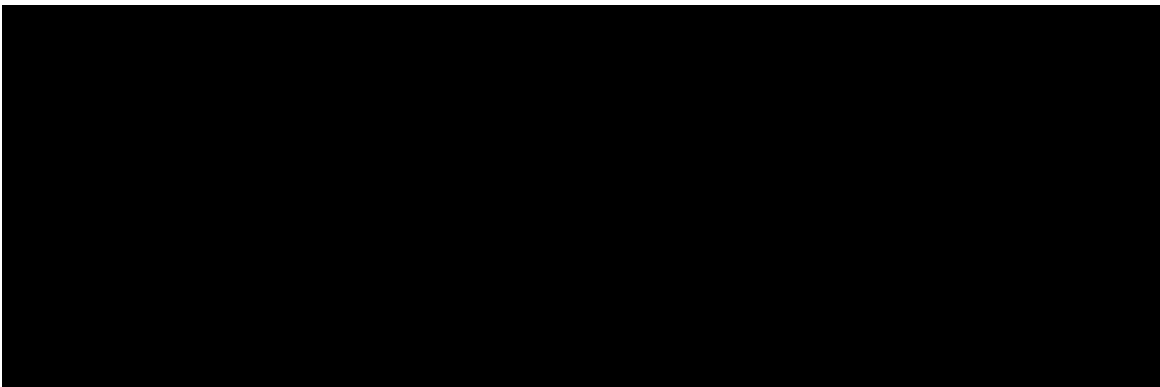
The PsAID-12 Questionnaire (Appendix N of Protocol) is a 12-item, self-administered questionnaire that reflects the impact of PsA from the perspective of the patient. It is composed of 12 physical and psychological domains. Each domain is rated from 0 to 10 with a different weighting. The total score is divided by 20.

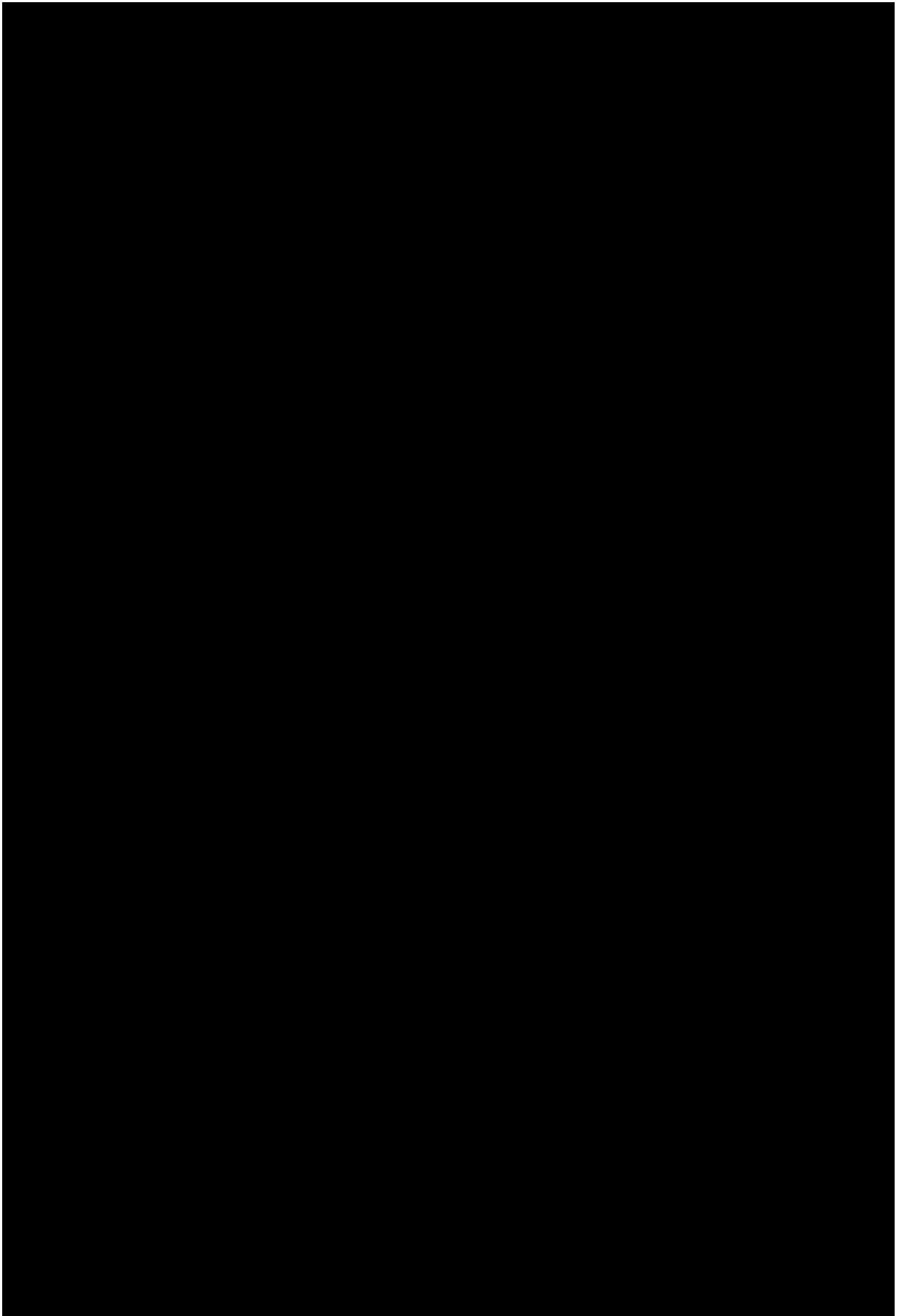
Final PsAID-12 value = [(PsAID1 (pain) NRS value (range 0-10) x 3)
+ (PsAID2 (fatigue) NRS value (range 0-10) x 2)
+ (PsAID3 (skin) NRS value (range 0-10) x 2)
+ (PsAID4 (work and/or leisure activities) NRS value (range 0-10)
x 2)
+ (PsAID5 (function) NRS value (range 0-10) x 2)
+ (PsAID6 (discomfort) NRS value (range 0-10) x 2)
+ (PsAID7 (sleep) NRS value (range 0-10) x 2)
+ (PsAID8 (coping) NRS value (range 0-10) x 1)
+ (PsAID9 (anxiety) NRS value (range 0-10) x 1)
+ (PsAID10 (embarrassment and/or shame) NRS value (range 0-
10) x1)
+ (PsAID11 (social life) NRS value (range 0-10) x 1)
+ (PsAID12 (depression) NRS value (range 0-10) x 1)]/20

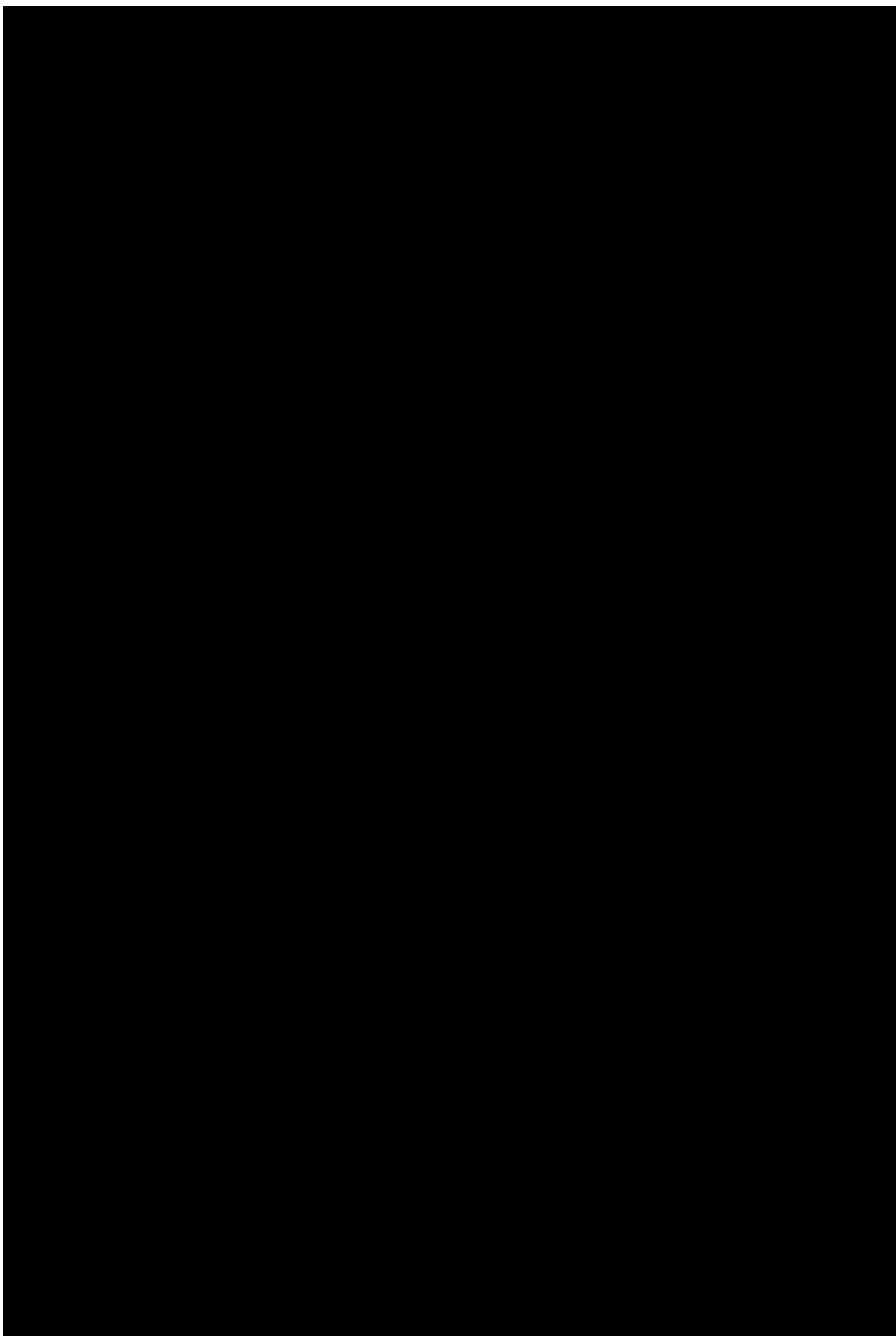
The final score has a range from 0 (best status) to 10 (worst status), with a cut-off ≤ 4 representing patient-acceptable symptom state. (Gossec, 2014).

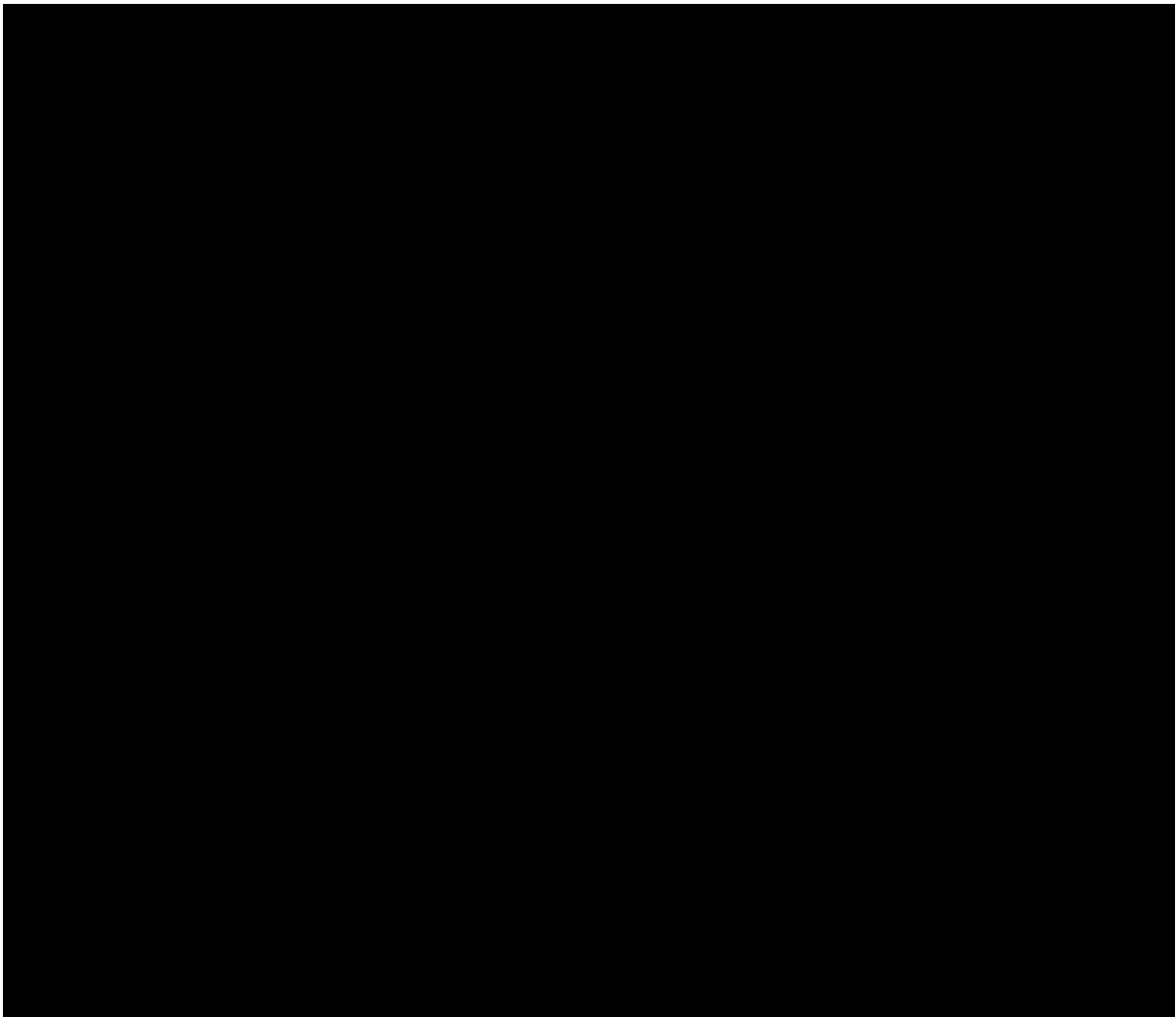
If one of the 12 domains composing the PsAID is missing, it will be imputed by the mean value of the 11 non-missing NRS scores. If two or more of the domains are missing, the final PsAID score will be considered as missing.

Data resource: PsAID-12 is captured via tablets provided by ERT.









5.2 Derivations of Safety Endpoints

Baseline definition for all safety endpoints is given in Section 5.3. Change from baseline is calculated as on-treatment value minus the baseline value. Handling of time points is described in Section 5.4.

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 investigational product (IP) and no later than 28 days after the last dose of IP for subjects who have completed the study 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 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 partially missing AE start dates are described in Section [8.3.1](#)

5.2.1 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 (action taken on AE eCRF is recorded as "DRUG WITHDRAWN"). 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.

5.3 Baseline Definitions

For both efficacy and safety analyses in the Placebo-controlled Phase, baseline will be 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 the last value measured prior to or on the first apremilast dose date.

5.4 Time Points

Time points in all efficacy analyses are based on the visits/study weeks as recorded in the database. 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 and discontinuation) measured or collected within the specific analysis phase/period 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. Only scheduled visit will be used for time point analysis.

- Early Termination Visit

A value from the Early Termination Visit (for a discontinued subject) will be mapped to the subsequent scheduled visit for that parameter after the last visit the subject has completed.

Note that each measurement collected at the Early Termination Visit will only be mapped to the next scheduled visit for that measurement. For example, SF-36 is scheduled to be

collected at Week 16/Visit 5, Week 24/Visit 6, and Week 48/Visit 9. If a subject discontinued after Week 24/Visit 6, then the SF-36 Early Termination data would be mapped to Week 48/Visit 9, not Visit 7.

Analysis visits in safety analyses are based on the remapped visits/study weeks using the visit mapping algorithm ([Appendix A](#)), which may or may not be the same as the visits/study weeks as recorded in the database.

5.5 Analysis Phases/Period

For the purposes of data analysis and reporting, below analysis phases/period for post-baseline data are defined: the Placebo-controlled Phases (through Week 24 and through Week 16) and the Apremilast-exposure Period.

The start and stop dates of each analysis phases/period are defined below, unless otherwise defined for a specific analysis or summary. Analyses for a specific period will include all data occurring, measured or collected (based on date of occurrence or measurement, or date of specimen collection, as appropriate) between the start and stop dates (inclusive or exclusive, depending on the data type) of that analysis phases/period. The specific rules for including or excluding the start and stop dates, depending on the data type, are also given below.

- Placebo-controlled Phase through Week 24:
 - Start date: The date of the first dose of IP.
 - Stop date:
 - The date of the first Week 16/Visit 5 dose for placebo-treated subjects who escape early at Week 16/Visit 5 and take at least one dose of IP dispensed at that visit, or
 - The date of the first Week 24/Visit 6 dose for subjects who complete Week 24/Visit 6, except for placebo-treated subjects who escape early, or
 - For all other subjects (i.e., subjects who discontinue prior to Week 24/Visit 6, with the exception of placebo-treated subjects who discontinue after early escape and has taken at least one dose of IP dispensed at Week 16/Visit 5), the date of the last study visit (including the Early Termination Visit and the Observational Follow-up Visit; i.e., all data after the start date will be included) for efficacy analyses, or 28 days after the last dose of IP or the date of death (if applicable), whichever comes first, for safety analyses.

- Placebo-controlled Phase through Week 16:
 - Start date: The date of the first dose of IP.
 - Stop date:
 - The date of the first Week 16/Visit 5 dose for placebo-treated subjects who escape early at Week 16/Visit 5 and take at least one dose of IP dispensed at that visit, or
 - The last non-missing Week 16/Visit 5 date for subjects who complete Week 16/Visit 5, except for placebo-treated subjects who escape early, or
 - For all other subjects (i.e., subjects who discontinue prior to Week 16/Visit 5), the date of the last study visit (including the Early Termination Visit and the Observational Follow-up Visit; i.e., all data after the start date will be included) for efficacy analyses, or 28 days after the last dose of IP or the date of death (if applicable), whichever comes first, for safety analyses.

For the Placebo-controlled Phases, the start date is exclusive, and the stop date is inclusive for efficacy analyses, and safety analyses of laboratory, vital signs, weight, and ECG data. For AE analyses, the start date is inclusive, the stop date under the first 2 bullets is exclusive and the stop date under the third bullet is inclusive.

- Apremilast-exposure Period:
 - Start date:
 - The date of the first dose of apremilast after being randomized (at Week 0/Visit 2) or escaped/transitioned (at Week 16/Visit 5 or Week 24/Visit 6) to an apremilast dose group.
 - Stop date:
 - For efficacy analyses, the date of the last study visit available in the database (including the Early Termination Visit and the Observational Follow-up Visit; i.e., all data after the start date of the Apremilast-exposure Period will be included), or
 - For safety analyses, 28 days after the last dose of apremilast or the date of death (if applicable), whichever comes first, for subjects who have completed the study or have discontinued early by the time of database cut. For ongoing subjects, the data cutoff

date for safety analyses of laboratory, vital signs, weight, and ECG data, and AE analyses.

For the Apremilast-exposure Period, the start date is exclusive and the stop date is inclusive for efficacy analyses, and safety analyses of laboratory, vital signs, weight, and ECG data. For AE analyses, both the start and stop dates are inclusive.

6. Analysis Sets

6.1 Full Analysis Set

The Full Analysis Set (FAS) will consist of all subjects who are randomized as specified in the protocol. Subjects who are randomized in error and do not receive any dose of IP will be excluded from FAS. Efficacy analyses for the Placebo-controlled Phase will be primarily based on the FAS and subjects will be included in the treatment group to which they are randomized.

6.2 Safety Analysis Set

The Safety Analysis Set will consist of all subjects who receive at least one dose of study medication. The safety analyses for the Placebo-controlled Phase will be based on the Safety Analysis Set and subjects will be included in the treatment group for the treatment they actually receive. For most subjects this will be the treatment group to which they are randomized. Subjects who take an IP that differs from the assigned one, according to the randomization schedule, for the entire Placebo-controlled Phase will be included in the treatment group corresponding to the IP they actually received. Subjects received at least one dose of apremilast in the Placebo-controlled Phase will have their actual treatment arm as apremilast.

6.3 Per Protocol Set(s)

The Per Protocol (PP) Set will consist of all subjects who are randomized and receive at least one dose of study medication and have no major protocol deviations during the Placebo-controlled Phase that will have a major impact on the assessment of the primary endpoint. The final determination of major protocol deviations will be made prior to the unblinding of treatment assignment and will be documented. The PP analyses in the Placebo-controlled Phase will be based on the PP population and subjects will be included in the treatment group for the treatment they are randomized.

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

Not applicable.

6.5 Pharmacokinetic/Pharmacodynamic Analyses Set(s)

Not applicable.

6.6 Interim Analyses Set(s)

Not applicable.

6.7 Study-specific Analysis Sets

6.7.1 Apremilast Subjects as Randomized or Transitioned Population

The efficacy analyses for the Apremilast-exposure period that include data from the placebo-controlled and active-treatment extension phases will be based on the subjects as randomized or transitioned population, which will include all FAS subjects who are randomized (at Week 0) or escaped/transitioned (at Week 16 or Week 24) to apremilast. The data summaries will be tabulated by the group subjects are randomized to (i.e., separately for subjects who are randomized to apremilast or placebo at Week 0).

6.7.2 Apremilast Subjects as Treated Population

Safety analyses for the Apremilast-exposure Period will be based on the apremilast subjects as treated, which will include all subjects who are treated with apremilast from Week 0 or transitioned to apremilast at Week 16 or Week 24. Subjects' safety data during the Apremilast-exposure Period will be included.

7. Planned Analyses

7.1 Interim Analysis and Early Stopping Guidelines

No interim analysis will be conducted.

7.2 Primary Analysis

After all subjects have completed the Week 24 Visit (or discontinued from the study), a Week 24 database restriction will be performed, the primary data analysis will be conducted. However, unblinded data will only be made available to selected Sponsor and Contract Research Organization (CRO) team members involved with analysis of the data. All other Sponsor, site, and CRO personnel directly involved with the conduct of the study, will remain blinded to treatment assignments until the final database lock at the conclusion of the study.

The primary analysis will present data from the completed placebo-controlled phase of the study, including the primary, secondary endpoints at Week 16, and [REDACTED] and safety results up to week 24.

7.3 Final Analysis

At the end of the study, after all subjects have completed or have been discontinued from the Apremilast Extension Phase (Weeks 24 to 48) and the Observational Follow-up Phase, the final analysis will be performed. The scope of the final analysis is the evaluation of the long-term efficacy and safety data. A final Clinical Study Report will be generated.

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 and will be processed by Amgen Global Statistical programming (GSP) to be used in the planned analyses. This study will use the RAVE database, the efficacy endpoint data received from Clario (formerly ERT), IWRS data from Endpoint and Laboratory data from LabCorp (formerly Covance).

8.3 Handling of Missing and Incomplete Data

8.3.1 Adverse Events

Partially missing AE start dates will be imputed in the ADaM dataset for AEs.

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

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 or Week 24 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: 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: 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: 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 24 or early escape date as “D_{APR}/M_{APR}/Y_{APR}”. The following table gives the imputation rules for partially missing AE start dates.

Table 3. 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 |
| 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 |
| 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}) |

8.3.2 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 with 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 with the latest possible date given the non-missing field(s) of the date.

8.3.3 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. If this imputation results in a medical treatment start date is on or after the informed consent date, then the incomplete medical treatment start date will be assigned as one day before earliest recorded informed consent date. If the end date is complete with no missing year, month, or day, and the partially missing start date imputed by the rule above is after the end date, then the start date will be imputed with the end date.

8.4 Detection of Bias

This study has been designed to minimize potential bias by the use of randomization of subjects into treatment groups and the use of blinding. Other factors that may bias the results of the study include:

- Major protocol deviations likely to impact the analysis and interpretation of the efficacy endpoints.
- Subject level unblinding before the primary analysis database lock and formal unblinding.

Important protocol deviations likely to impact the analysis and interpretation of the efficacy endpoints will be tabulated in the Clinical Study Report (CSR).

Any unblinding of individual subjects prior to formal unblinding of the study will be documented in the CSR. The impact of such unblinding on the results observed will be assessed.

Additional sensitivity analyses will be performed to assess the impact of potential biases on the primary endpoint if needed. If any sensitivity analyses are required to evaluate

potential biases in the study's conclusions, then the sources of the potential biases and results of the sensitivity analyses will be documented in the CSR.

8.5 Outliers

Extreme data points will be identified during the blinded review of the data prior to primary analysis database lock. Such data points will be reviewed with clinical data management to ensure accuracy. The primary analyses will include outliers in the data. Sensitivity analyses that exclude outliers may be undertaken if extreme outliers for a variable are observed.

8.6 Distributional Characteristics

Distributional assumptions for the primary and secondary endpoints will be assessed. If the assumptions are not met, then alternative methods will be utilized. The use of alternative methods will be fully justified in the CSR.

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

9.1.1 Reporting and Analysis Conventions

Summary statistics for continuous variables include sample size (n), mean, standard deviation (SD), median, minimum, the 25th (Q1) and 75th (Q3) percentiles, and maximum. All mean and median values will be formatted to one more decimal place than the measured value. SD 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. Frequency summary for categorical variables includes number and percentage. All percentages will be rounded to one decimal place, unless presenting two decimal places is necessary when the percentages are very small. Number and percentage values will be presented as xx (xx.x%). All analysis and summary tables will have the population sample size for each treatment group in the

column heading. P-values (2 sided) will be presented with 4 decimal places. All laboratory data will be reported using standard international (SI) units.

Change from baseline will be calculated as (post-baseline minus baseline value).

Percent (%) change from baseline will be calculated as (post-baseline minus baseline value) divided by (baseline value), then multiply by 100. If a baseline value is zero, the percent change from baseline value will be excluded from the summaries.

The analyses involving stratification factors (use of glucocorticosteroid at baseline and use of a csDMARD) will be primarily based on the stratification factors recorded in IWRS at the randomization. Similarly, EE status, as recorded in the IWRS, will be used in all derivations/analyses.

Subject data listings will be provided to support the tables and graphs for selected parameters.

9.1.2 Statistical Methodology for Efficacy

For the Placebo-controlled Phase, treatment outcomes between APR 30 mg BID and placebo will be compared primarily using the FAS. Supportive analysis of the primary endpoint (MDA-joints at Week 16) will use the PP.

For efficacy analyses, data obtained after the day of the first Week 16/Visit 5 dose for subjects who escape early at Week 16/Visit 5 and take at least one dose of IP dispensed at that visit will be excluded from the placebo-controlled summaries through Week 24/Visit 6, unless otherwise specified.

The null and alternative hypotheses for the treatment comparisons can be described as follows:

$$H_0: p_{apremilast} = p_{placebo} \quad \text{vs.} \quad H_1: p_{apremilast} \neq p_{placebo}$$

Where p stands for the parameter in comparison, e.g., proportion of subjects achieving MDA-Joints.

Efficacy results will be considered statistically significant after incorporating the strategy for controlling the family-wise Type 1 error rate as described in Section 9.1.3, Multiplicity. All statistical tests will be conducted at the $\alpha = 0.05$ (2-sided) level. That is, 2-sided p-values and 2-sided confidence intervals (CIs) for intended point estimates will be reported.

Analyses of the long-term data that include the Placebo-Controlled Phase and the Active-treatment Extension Phase will be based on Apremilast Subjects as Randomized population.

9.1.3 Multiplicity

Among statistical tests for the comparisons between the apremilast 30 mg BID and placebo groups, the multiplicity of the analyses performed for the primary and secondary efficacy endpoints will be adjusted using a sequential test procedure to preserve the family-wise type I error rate. The secondary efficacy endpoints and the hierarchical order of the endpoints in the sequential testing will follow those listed in Section [2.1](#)

The sequential test procedure will be conducted as follows:

- Formal statistical tests will be carried out sequentially for selected endpoints, starting with the primary endpoint and then the secondary endpoints by the order same as presented in Section [2.1](#)
- At each step in the sequential testing, statistical significance for a specific endpoint will be claimed only if the p-value for the test is less than 0.05, and the p-values for the tests of all previous endpoints in the hierarchy order are also less than 0.05.
- Once a p-value is 0.05 or greater for the testing of a specific endpoint, only the tests for the preceding endpoints will be considered statistically significant; the testing for this and the subsequent endpoints will not be considered statistical significance and the nominal p-values will be reported.

Of note, the above multiplicity control will only be conducted for the primary endpoint and the secondary endpoints that are evaluated in the Placebo-controlled Phase.

9.2 Subject Accountability

The number of subjects screened, randomized, receiving IP, and completing the study will be summarized.

Subject disposition will be tabulated by treatment arm and study phase. Reasons for discontinuation will be summarized.

The number and percent of subjects randomized will be tabulated by the stratification factors.

9.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's initial visit and updated during the IPD reviews throughout the study prior to primary analysis database lock. These definitions of IPD categories, subcategory codes, and descriptions will be used during the course of the study. Eligibility deviations are defined in the protocol.

The protocol deviations/violations will be identified and assessed by clinical research physician or designee throughout all study periods. The protocol deviations/violations for a particular analysis will be defined prior to unblinding of the database. For example, the analysis for the primary endpoint will be based on the 24-week database and definitions of major protocol deviations/violations that may affect the analysis will be determined prior to unblinding of the 24-week database. The protocol deviations/violations will be summarized by treatment arm for the FAS for the double-blind, Placebo-controlled Phase (up through Week 24, regardless of early escape status). Additional summaries will be provided based on Apremilast Subjects as Randomized through Week 52 (Weeks 0-52).

A list of protocol deviations/violations for all subjects will be provided by the Clinical Research Physician and/or through a standardized algorithm prior to the primary analysis database lock and unblinding. Among these protocol deviations, those that may have impact on treatment efficacy will also be determined prior to the primary analysis database lock and used to identify which subjects are to be removed from the Per Protocol Set. These protocol deviation data will be integrated into the clinical database. Additionally, protocol deviations and important protocol deviations related to COVID-19 will be summarized.

9.4 Demographic and Baseline Characteristics

Demographics and baseline characteristics will be summarized descriptively by treatment group using the FAS. In addition, demographics and baseline disease characteristics will also be summarized for subjects who escape early, subjects who do not escape early and subjects who discontinue early. The comparability of the treatment groups for each relevant characteristic will be assessed descriptively in table format; no statistical hypothesis tests will be performed on these characteristics.

9.4.1 Demographics

Demographic characteristics including age, sex, race, ethnicity, geographic region, weight (kg) and height (cm), etc. (please refer to CRF for available variables) will be summarized.

Continuous variables will be summarized using descriptive statistics (n, mean, standard deviation, median, minimum, and maximum, etc.). Categorical variables will be summarized with counts and percentages.

Summary statistics will be provided by treatment group for the following continuous variables:

- Age (years)
- Weight (kg)
- Height (cm)
- [REDACTED]

Number and percentage will be provided by treatment group for the following categorical variables where sufficient numbers allow:

- Age category (< 65, ≥ 65 years; < 40, ≥ 40 to < 65, ≥ 65 to < 75, ≥ 75 to < 85, ≥ 85 years, if appropriate)
- Sex (male, female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islanders, White, Not Collected or Unknown)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown)
- Region (North America, Europe, Rest of the World [including Russia])
- Weight category (< 70, ≥ 70 to < 85, ≥ 85 to < 100, ≥ 100 kg)
- [REDACTED]

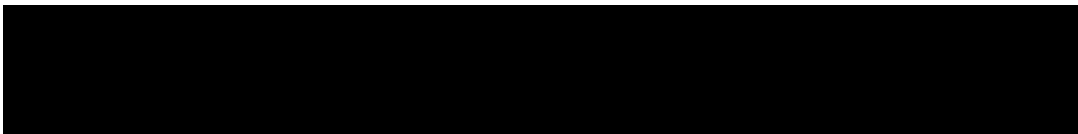
Most variables will be used as recorded on the CRF.

The demographics summary will include the frequencies of alcohol history (current user, past user, never used), and smoking status (current user, past user, never used).

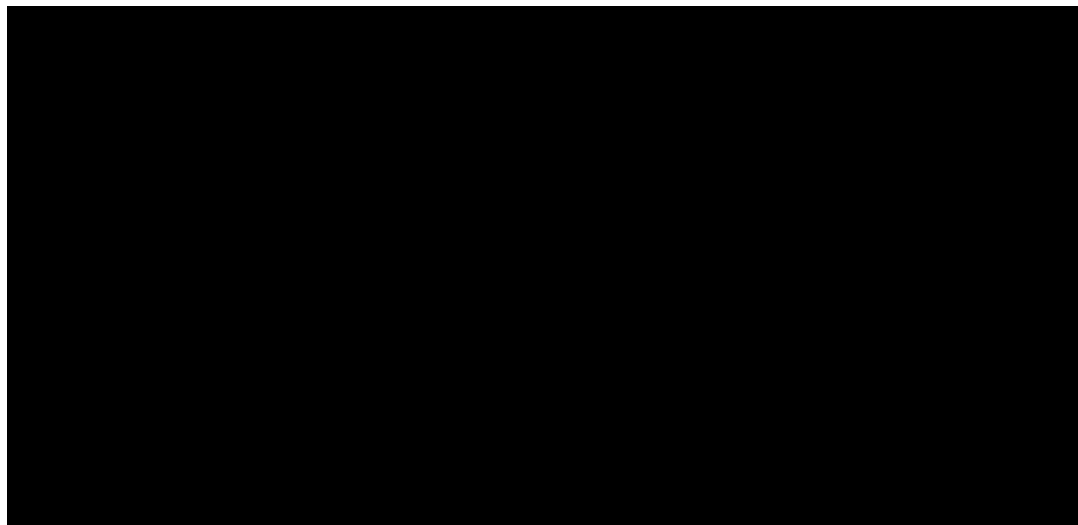
9.4.2 Baseline or Disease Characteristics

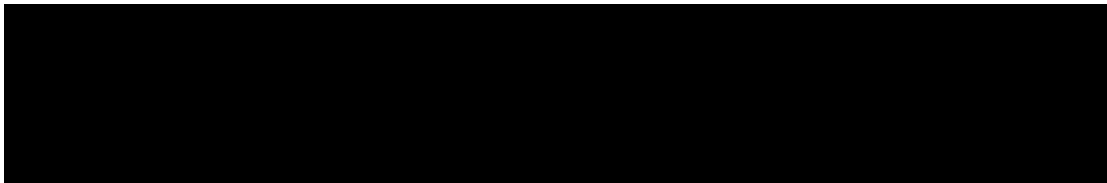
Summary statistics will be provided by treatment group for all relevant baseline characteristics continuous variables, such as the following:

- Duration of PsA (from the date of diagnosis to the date of informed consent; months)
- Baseline prednisone (or equivalent) dose (mg/day)
- Concomitant Methotrexate dose (mg/week)
- Concomitant Sulfasalazine does (g/day)
- cDAPSA
- Swollen joint count (66 joints)
- Tender joint count (68 joints)
- Patient's global assessment of disease activity (0 to 100 mm VAS)
- Patient's assessment of pain (0 to 100 mm VAS)
- PsAID-12
- HAQ-DI score and categories
- Physician's global assessment of disease activity (0 to 100 mm VAS)
- LEI in subjects with pre-existing enthesopathy



- PASDAS
- Psoriasis BSA

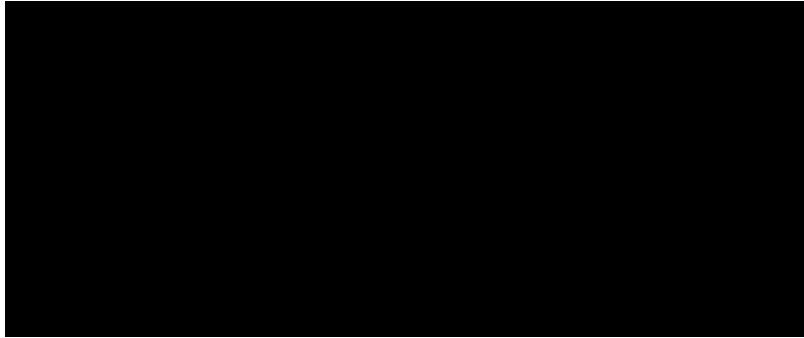




- hsCRP (mg/L)

Number and percentage will be provided by treatment group for all relevant categorical variables, such as the following (subcategories under a specific category of a variable, where applicable):

- Baseline glucocorticosteroid use (yes, no) in IWRS and in actual data, if different
- Prior/concomitant use of a csDMARD, either MTX or SSZ, (csDMARD-naïve; csDMARD use prior to baseline; or csDMARD use prior to and concomitant at baseline) in IWRS and in actual data, if different
- Prior use csDMARD (yes, no)
- Concomitant use of csDMARD (yes, no)
- Tender joint count (2, 3, 4)
- Swollen joint count (2, 3, 4)
- Swollen or tender joint count (2, 3, 4, 5, 6, 7, 8)
- Swollen and tender joint count (0, 1, 2, 3, 4)
- cDAPSA (≤ 4 for remission, > 4 and ≤ 13 for low disease activity, > 13 and ≤ 27 for moderate disease activity, and > 27 for high disease activity)
- Patient's global assessment of disease activity VAS (≤ 20 , > 20)
- Patient's assessment of pain VAS (≤ 15 , > 15)
- PsAID-12 (≤ 4 , > 4)
- PASDAS (≤ 3.2 , > 3.2 to < 5.4 , ≥ 5.4)
- HAQ-DI (≤ 0.5 , > 0.5)
- LEI (= 0, > 0)
- LEI (≤ 1 , > 1)
- Psoriasis BSA ($\leq 3\%$, $> 3\%$)



9.4.3 Medical History

Medical history will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA, version 25.1 or later). A frequency summary (number and percentage) of medical history will be presented by treatment group, system organ class (SOC), and preferred term (PT).

9.4.4 Prior and Concomitant Procedures

The prior and concomitant procedures will be coded using the MedDRA, version 25.1 or higher. Frequency summaries of prior procedures for Full Analysis Set, and concomitant procedures for Safety Analysis Set will be provided by treatment group, SOC and PT.

Prior procedures are defined as procedures or surgeries that were started before the first dose of IP (regardless of they ended before the first dose of IP). Prior procedures that continue or occur after the first dose of IP will also be reported as concomitant procedures.

Concomitant procedures will be similarly analyzed as concomitant medications in the next section.

9.4.5 Prior and Concomitant Medications

The Anatomical Therapeutic Chemical (ATC) coding scheme of the World Health Organization Drug Dictionary (WHODD version March 2020 or later) will be used to group medications into relevant categories. Frequency summaries of prior medications for Full Analysis Set and concomitant medications for Safety Analysis Set will be provided by treatment group, ATC2 level, and standardized medication name.

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 be also reported as concomitant medications.

Concomitant medications and procedures will be summarized differently for each phase/period of the study. The concomitant medication for the Placebo-controlled Phase will be summarized by treatment group using the Safety Analysis Set. The concomitant medication for the Apremilast-exposure Period will be summarized by a total group only of those who were treated with apremilast (irrespective of when the apremilast exposure starts [at Week 0, 16, or 24]). In addition, to account for possible resumption/initiation of non-study medications treating PsA during the Observational Follow-up Phase, concomitant medications will be summarized for the Observational Follow-up Phase based on all subjects who entered the phase.

A concomitant medication for the Placebo-controlled Phase is defined as a non-study medication that meets either one of the following 2 criteria:

1. Is started on or after the date of the first dose of IP, and
 - before the date of the first Week 16/Visit 5 dose for placebo-treated subjects who escape early at Week 16/Visit 5 and take at least one dose of IP dispensed at that visit, or
 - before the date of the first Week 24/Visit 6 dose for subjects who complete Week 24/Visit 6 dose, except for placebo-treated subjects who escape early, or
 - on or before the date of the last study visit (including the Early Termination Visit and excluding any unscheduled visits associated with the Early Termination Visit and the Observational Follow-up Visit) for all other subjects (i.e., subjects who discontinue prior to Week 24/Visit 6, with the exception of placebo-treated subjects who discontinue after early escape and have taken at least one dose of IP dispensed at Week 16/Visit 6).
2. Is started before the date of the first dose of IP, and stopped on or after the date of the first dose of IP, or recorded as ongoing.

A concomitant medication for the Apremilast-exposure Period is defined as a non-study medication that meets either one of the following 2 criteria:

1. is started on or after the date of the first dose of apremilast after being randomized (at Week 0/Visit 2) or escaped/transitioned (at Week 16/Visit 5 or Week 24/Visit 6) to apremilast dose group and stopped on or before the date of the last study visit (including the Early Termination Visit and excluding any unscheduled

visits associated with the Week 48/Visit 9 and the Early Termination Visit, and the Observational Follow-up Visit) for subjects who have completed the study or have discontinued early, or on or before the data cutoff date for ongoing subjects.

2. is started before the date of the first dose of apremilast after being randomized or escaped/transitioned to apremilast dose, and stopped on or after the date of the first dose of apremilast, or recorded as ongoing.

A concomitant medication for the Observational Follow-up Phase is defined as a non-study medication that meets either one of the following 2 criteria:

1. is started after the date of the Week 48/Visit 9 (for subjects who have completed the study) or the Early Termination Visit (for subjects who have discontinued early), but no later than 28 days after the last dose of IP in the study.
2. is started on or before the date of the Week 48/Visit 9 (for subjects who have completed the study) or the Early Termination Visit (for subjects who have discontinued early), and stopped after this date, or recorded as ongoing.

9.5 Efficacy Analyses

9.5.1 Analyses of Primary Efficacy Endpoint

The primary efficacy endpoint, proportion of subjects who achieve the modified MDA (MDA-Joints) at Week 16, will be compared between the Apremilast 30 mg BID group and placebo group using the Cochran-Mantel-Haenszel (CMH) test, controlling for use of glucocorticosteroids at baseline (yes/no) and use of a csDMARD (naïve, prior use only, or prior and concomitant use). Unadjusted group responder rates, unadjusted treatment differences in proportions, and adjusted treatment differences in proportions using the weighted average of the treatment differences across the strata with the CMH weights, along with the associated 2-sided 95% CIs using a normal approximation to the weighted average and the 2-sided p-values from the CMH tests, will be provided. Apremilast will be declared efficacious if the test result for the comparison is statistically significant in favor of Apremilast.

Subjects who have discontinued prior to Week 16 due to lack of efficacy or adverse event will be treated as non-responders in the analysis; and the multiple imputation methodology will be utilized, as appropriate, to impute the data for subjects whose

modified MDA at Week 16 is missing for a reason other than lack of efficacy or adverse event.

The combination of non-responder imputation of some subjects and the multiple imputation of the rest of missing values at Week 16 is to address the Missing Not At Random (MNAR) nature of missing values at the analysis time point. The detailed data handling and analysis follow the procedures below:

Step 1. Censor or apply NRI (non-responder imputation) per different scenario

- (Censoring) After applying Early Termination Visit data, missing Week 16 values remain as missing for patients discontinued prior to Week 16 due to other reasons
- (NRI) Set non-responder (NR) to subjects discontinued prior to Week 16 due to
 - Lack of efficacy
 - Adverse event

Step 2. Implement multiple imputation, with 40 iterations (using fully conditional specification to impute the missing modified MDA generated from Step 1).

Step 3. Calculate statistics on each imputation iteration from Step 2.

Step 4. Combine results from Step 3.

The overall p-value will be derived from the CMH test that is normalized *via* the Wilson-Hilferty transformation (Wilson & Hilferty, 1931; Goris, 1992).

Sensitivity analysis

Subjects with no response data to determine the response status at all assessment timepoints will be excluded.

The effect of protocol compliance will be studied by repeating the primary analysis using the PP population, according to the treatment subjects actually receive and the stratification factors subjects' medication data show.

In order to investigate the impact of the imputation method used for the primary analysis, a sensitivity analysis will be performed using the non-responder imputation for subjects who have discontinued prior to Week 16, or do not have sufficient information for a definitive determination of the modified MDA (MDA-Joints) status at Week 16. The responder analysis using the CMH model as in the primary analysis will be repeated.

To examine the robustness of the results obtained from the primary analysis based on the CMH method, a sensitivity analysis will be implemented using logistic regression. The responder analysis based on the logistic regression model will include treatment group, concomitant use of glucocorticosteroids at baseline (yes/no), and use of csDMARD (naïve, prior use only, or prior and concomitant use) as factors. This analysis will be based on the FAS data set used for the primary analysis with non-responder imputation along with the MI.

9.5.2 Analyses of Secondary Efficacy Endpoint(s)

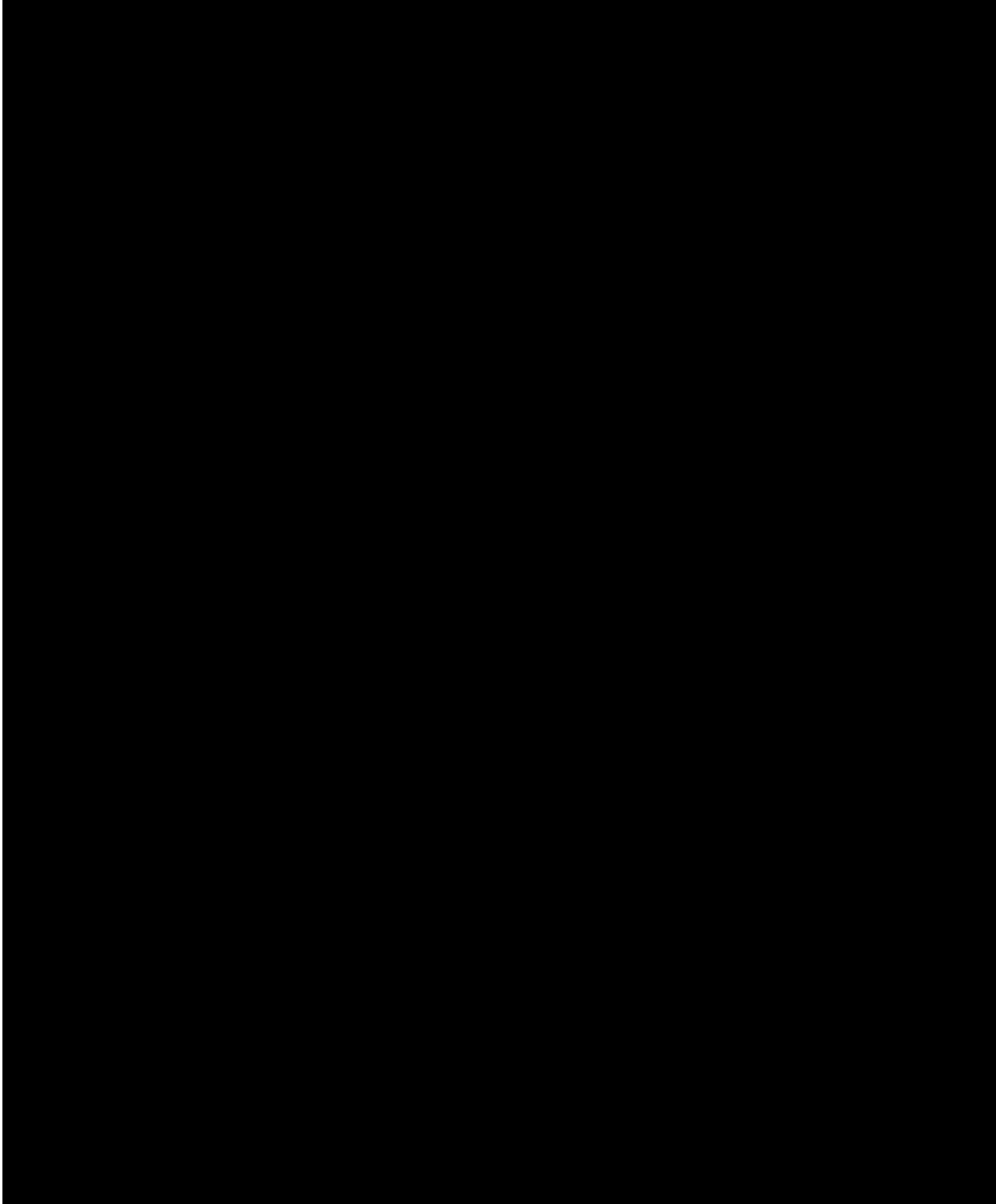
For the secondary endpoints, the missing data of the binary response parameters will be imputed using the non-responder imputation (NRI) and MI, as appropriate. The binary responses in the Placebo-controlled Phase will be analyzed using the CMH method, adjusted for use of glucocorticosteroids at baseline (yes/no) and use of a csDMARD (naïve, prior use only, or prior and concomitant use).

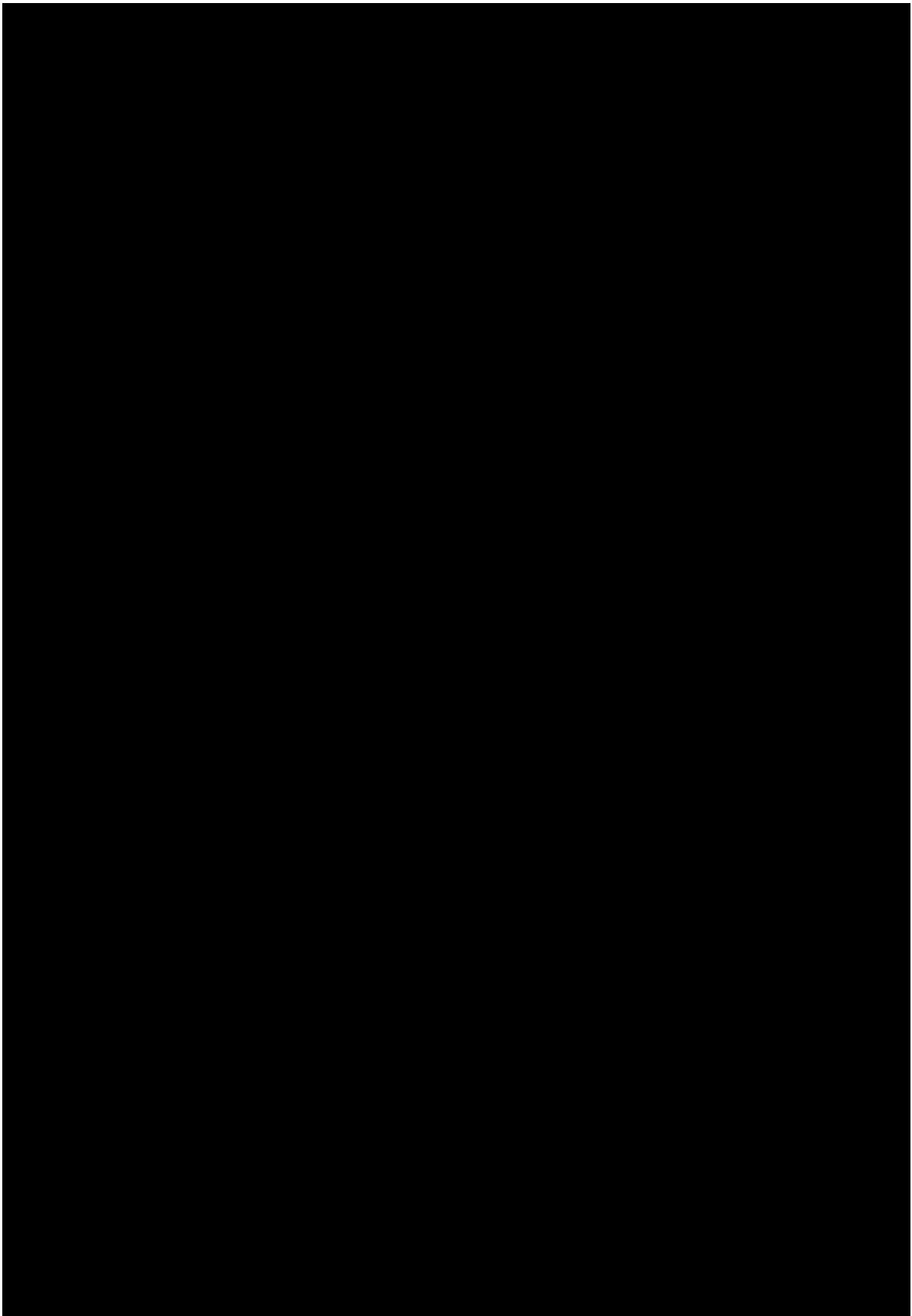
A mixed-effect model of repeated measures (MMRM) will be used to analyze the continuous endpoints. Because of the nature of the MMRM method, data from all timepoints during the Placebo-controlled Phase will be included in the model. Data obtained after early escape for subjects who escape early at Week 16 will be set to missing. Change from baseline will be analyzed using the MMRM approach with treatment group, time, treatment-by-time interaction, use of glucocorticosteroids (yes/no), use of csDMARD (naïve, prior use only, or prior and concomitant use) as factors, and baseline value as a covariate. Time will be treated as a categorical variable in the MMRM model. An unstructured covariance matrix will be used to model the within subject covariance. If the MMRM model using an unstructured covariance matrix fails to converge with the default Newton-Raphson algorithm, the Fisher scoring algorithm or other appropriate methods can be used to provide initial values of the covariance parameters. In the rare event that none of the above methods yields convergence, a structured covariance such as heterogeneous Toeplitz and Toeplitz structures will be used to model the correlation among repeated measurements. In this case, the empirical option will be used because the sandwich variance-covariance estimator is asymptotically consistent. Within-group least-squares (LS) mean changes from baseline, the associated standard errors (SEs) and 2-sided 95% CIs, treatment differences in LS mean changes from baseline at Week 16 and the associated 2-sided 95% CIs and 2-

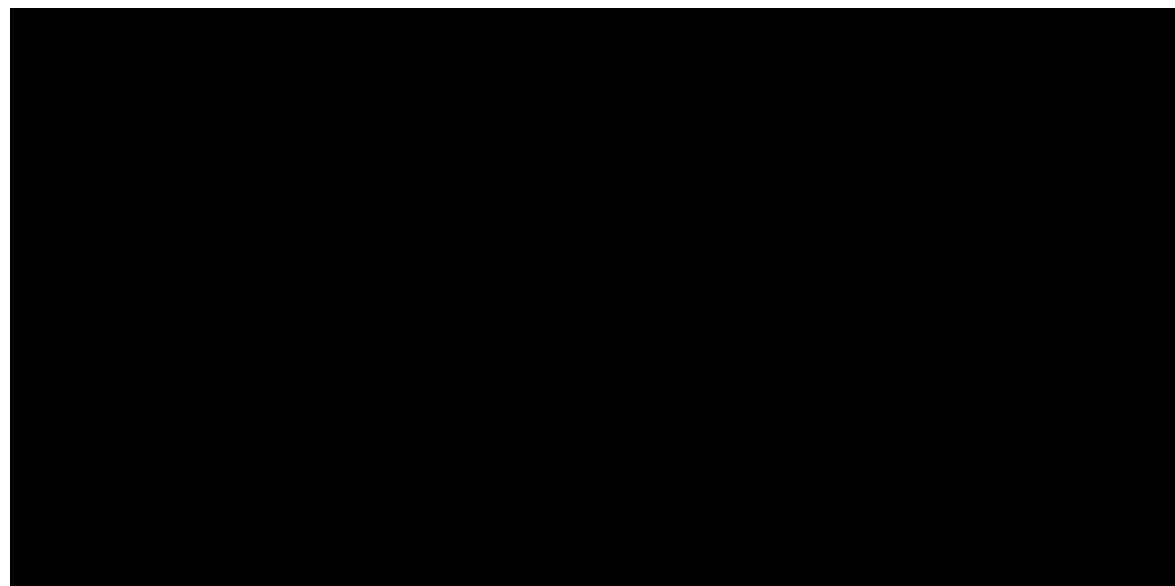
sided p-values will be obtained from the MMRM model and will be provided in the summary table.

All analyses on endpoints including joint counts will be performed on those joints that are affected at baseline (defined as sentinel joints).

9.5.3 Analyses of Exploratory Efficacy Endpoint(s)







9.6 Safety Analyses

9.6.1 Analyses of Primary Safety Endpoint

Safety will be evaluated via descriptive statistics and point estimates. Unless otherwise specified, all safety analyses described in this section will be performed for both the Placebo-controlled and Apremilast-exposure Phases (Section 5.5).

Safety analyses for the Placebo-controlled Phase will be based on the safety population and presented by treatment group (apremilast 30 mg BID and placebo); the data included are up to Week 16/Visit 5 for placebo-treated subjects who early escape and up to Week 24/Visit 6 for all other subjects. For the Apremilast-exposure Period, the analyses will be based on the apremilast subjects as treated population and include data from Day 1 of apremilast exposure, irrespective of when the apremilast exposure starts (at Week 0, 16, or 24). For both the Placebo-controlled Phase and the Apremilast-exposure Period, the data included will be up to 28 days after the last dose of IP or the date of death (if applicable), whichever comes first, for subjects who discontinue prematurely.

For the analyses of AEs and marked abnormalities, the following point estimates are provided, unless otherwise specified:

- 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, multiplied by 100. Subjects with multiple occurrences of a specific event in a defined exposure interval 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 a defined exposure interval will be counted only once in the numerator. 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.
 - Total exposure time: 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 days. 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. If the subject experiences an event which starts on the first dose of the phase/period, the exposure time of that subject for that event is 1 day.

In summaries of continuous variables with summary statistics by time point, values for baseline, time point, and change from baseline will be summarized at each time point for subjects who have values at both baseline and the specific post-baseline time point. Frequency summaries of shifts from baseline to post-baseline time points will be provided for subjects who have values at both baseline and each individual specific post-baseline time point. In addition to the scheduled visits, shift summaries will also be provided from baseline to the end of phase/period and to the worst value, including scheduled and unscheduled visits, in a specific post-baseline phase/period.

9.6.2 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 25.1 or later will be used to code all events categorized as adverse events to a system organ class and a preferred term with SOC presented in the standard international order and PTs within SOC will be presented in descending order of subject incidence.

The subject incidence of adverse events will be summarized for all treatment-emergent adverse events, drug-related treatment-emergent adverse events, severe adverse events, serious adverse events, adverse events leading to discontinuation of investigational product, fatal adverse events and adverse events of interest.

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

An overall summary of the following AE categories will be provided:

- 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

Summarization of the following AE categories will be provided:

- Any TEAE (by SOC and PT, and by PT only)
- Common TEAE (subject incidence $\geq 5\%$ by SOC and PT, and by PT only)
- Any serious TEAE (by SOC and PT)
- Any non-serious TEAE (by SOC and PT)
- Any drug-related TEAE (by SOC and PT)
- Any serious drug-related TEAE (by SOC and PT)
- TEAEs by maximum severity (by SOC and PT)
- Any TEAE leading to study drug withdrawal (by SOC and PT)
- Any TEAE leading to study drug interruption (by SOC and PT)
- Any TEAE leading to death (by SOC and PT), if applicable

If a subject reports multiple occurrences of a specific event within a specific analysis phase/period, the subject will be counted only once by the maximum severity (mild, moderate, severe, and, if needed, missing). 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 occurrences, the subject will be counted only once in the “missing” category of severity.

The relationship of an AE to IP is recorded on the AE CRF page as “not suspected” and “suspected”. Adverse events recorded as “suspected” or with missing relationship to IP will be assessed as drug-related.

Summaries about selected TEAEs will also be provided by onset day and the event duration.

New events of all TEAEs by exposure interval will be summarized for the Placebo-controlled Phase (≤ 1 , > 1 to ≤ 4 , > 4 to ≤ 8 , > 8 to ≤ 12 , > 12 weeks) and the Apremilast-exposure Period (≤ 1 , > 1 to ≤ 4 , > 4 to ≤ 8 , > 8 to ≤ 12 , > 12 to ≤ 16 , > 16 to ≤ 28 , > 28 to ≤ 40 , > 40 weeks). An event with a start date falling into an exposure interval is considered a new event for that interval. Each subject is counted once in the numerator of a subject incidence 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.

All TEAEs will be summarized by age, sex, glucocorticosteroid use at baseline and prior/concomitant use of csDMARD. Subgroups of glucocorticosteroid use at baseline and prior/concomitant use of csDMARD will be based on the actual data, not by the data from IWRS.

If there are deaths, a listing will be provided.

9.6.3 Adverse Events of Special Interest

AESIs include protocol defined diarrhea AEs, which refers to events with two or more watery or liquid stools in a day, will be captured on the tablet by ERT. Frequency of the individual diarrhea will be captured.

Frequencies (Daily, 5 - 6 days per week, 2 - 4 days per week, once every other week to once a week, once a month or less) of these events will be summarized.

The summaries will be generated for the Placebo-controlled Phase, and for the Apremilast-exposure Period.

9.6.4 Laboratory Test Results

Summary statistics of observed values and changes from baseline in laboratory parameters will be provided by time points. For each hematology and chemistry analyte, frequency summaries of shifts from baseline to post-baseline time points and to the worst post-baseline value during each treatment phase will be provided. The low,

normal, and high categories of laboratory values used in the shift will be based on the laboratory normal range.

Laboratory marked abnormalities will be summarized based on the criteria in marked abnormalities criteria ([Appendix B](#)). 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 separate summary of laboratory marked abnormalities will also be presented by categories of subjects with normal values at baseline and subjects with abnormal values at baseline. To the purpose 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. A subject data listing of laboratory marked abnormalities will be provided. In the listing, all recordings of a lab analyte will be displayed for the subject if that subject has one or more marked abnormalities for the test.

For comorbid conditions of interest or a subgroup of subjects, summary statistics of relevant laboratory parameters over time may be provided. Subject data listings of all laboratory data, including urinalysis, will be provided.

Clinical laboratory evaluations will be available from the central laboratory and include:

- Hematology: complete blood count (red blood cell [RBC] count, hemoglobin, hematocrit, white blood cell [WBC] count and differential, absolute WBC counts, platelet count).
- Serum chemistries: total protein, albumin, calcium, phosphorous, glucose, triglycerides, total cholesterol [TC], total bilirubin, alkaline phosphatase, aspartate aminotransferase [AST; serum glutamic-oxaloacetic transaminase, SGOT], alanine aminotransferase [ALT; serum glutamic-pyruvic transaminase, SGPT], sodium, potassium, chloride, bicarbonate [carbon dioxide, CO₂], blood urea nitrogen, creatinine, creatinine clearance [estimated by the Cockcroft–Gault equation], lactate dehydrogenase [LDH], magnesium
- Dipstick urinalysis: specific gravity, pH, glucose, ketones, protein, blood, bilirubin, leukocyte esterase, nitrite, and urobilinogen. Dipstick urinalysis will be performed by the central laboratory; if the dipstick urinalysis is abnormal then a microscopic urinalysis (epithelial cells, RBC, WBC, and casts) will be performed.

A subject listing will be provided for all available pregnancy test results.

9.6.5 Vital Signs

For vital signs (temperature, pulse, and blood pressure) and weight, summary statistics (n, mean, standard deviation, median, minimum, the 25th (Q1) and 75th (Q3) percentiles, and maximum) of observed and change from baseline values will be presented. For vital signs (excluding weight), shift from baseline to each post-baseline visit, to worst (low/normal/high/both low and high) and to end of phase/period in below, within and above the normal ranges will be displayed in cross-tabulations for the Placebo-controlled Phase and for the Apremilast-exposure Period for parameters included in below Table 4. For weight, change from baseline (and percent change from baseline) to end of phase/period will be presented, as well as by baseline [REDACTED] and baseline weight categories.

Table 4. Normal Ranges for Shift Table

| Vital Sign Parameter | Normal Ranges |
|--------------------------|------------------------|
| Pulse | 60 – 100 beats/minutes |
| Systolic blood pressure | 90 – 140 mmHg |
| Diastolic blood pressure | 60 – 90 mmHg |

9.6.6 Physical Measurements

Physical examination findings will be captured as medical history at screening. Abnormal physical examination findings post Screening will be reported as adverse events and will be reported in the summaries.

9.6.7 Electrocardiogram

Not applicable.

9.6.8 Antibody Formation

Not applicable.

9.6.9 Exposure to Investigational Product

Descriptive statistics will be produced to describe the exposure to investigational product by treatment group. Compliance to each IP will also be summarized.

Treatment duration will be summarized by treatment group for the Placebo-controlled Phase using the safety population.

Summary statistics for treatment duration (in weeks), as well as a frequency summary of treatment duration categories (e.g., < 4, ≥ 4 - < 8 weeks, etc.) and exposure categories (e.g., ≥ 1 day, ≥ 4 , ≥ 8 weeks, etc.), will be provided.

Treatment duration (in weeks) is calculated as (the date of the last dose of IP – the date of the first dose of IP + 1) / 7 and rounded to one decimal place.

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 the last dose date is completely missing, set the last dose date to the minimum of (the 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, clinical efficacy assessment date, lab, vital signs, AE start or end dates, concomitant medications start or end dates, concomitant procedure date, treatment exposure start date or end date where doses were completely or partially taken, death date).

The specific definitions of the first and last dose dates for the Placebo-controlled Phase and Apremilast-exposure Period are given below:

- Placebo-controlled Phase:
 - First dose date: The date of the first dose of IP.
 - Last dose date:
 - the date prior to Week 16/Visit 5 for placebo-treated subjects who escape early at Week 16/Visit 5 and take at least one dose of IP dispensed at that visit, or
 - the date prior to Week 24/Visit 6 for subjects who complete Week 24/Visit 6 dose, except for placebo-treated subjects who escape early
 - For all other subjects (i.e., subjects who discontinue prior to Week 24/Visit 6), the date of the last dose of IP in the study.

- Apremilast-exposure Period:
 - First dose date:
 - The date of the first dose of apremilast after being randomized (at Week 0/Visit 2) or transitioned (at Week 16/Visit 5 or Week 24/Visit 6) to apremilast.
 - Last dose date:
 - The date of the last dose of apremilast.

The treatment compliance (in %) 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% - ≤120%, and > 120%.

9.6.10 Exposure to Non-investigational Product

Not applicable.

9.6.11 Exposure to Other Protocol-required Therapy

Not applicable.

9.7 Other Analyses

Not applicable

9.7.1 Analyses of Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Endpoints

Not applicable.

9.7.2 Analyses of Clinical Outcome Assessments

Not applicable.

9.7.3 Analyses of Health Economic Endpoints

Not applicable.

9.7.4 Analyses of Biomarker Endpoints

Not applicable.

10. Changes From Protocol-specified Analyses

During the development of this document, it was realized that the analysis for endpoint

specified in section 2, Table 12 of the protocol is a typing error and that

is desired. The protocol has not been amended for

this. The changes will be documented in the Clinical Study Report.

11. Literature Citations / References

Goria, M.N. 1992. On the fourth root transformation of chi-square. Australian Journal of Statistics, 34 (1), 55-64.

Rubin D.B. 1987. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons.

Ware J E., Kosinski M, Keller S D, 1994. SF-36 physical and mental health summary scales: a user's manual, Boston : Health Institute, New England Medical Center.

Wilson, E.B., Hilferty, M.M. 1931. The distribution of chi-squared. Proceedings of the National Academy of Sciences, Washington, 17, 684–688.

12. Prioritization of Analyses

Not applicable.

13. Data Not Covered by This Plan

Not applicable.

14. Appendices

Appendix A. Analytical Windows

For safety analyses, all available values obtained between 2 planned visits including unscheduled measurements will be pooled to the analysis visit window according to the following definition. In the event of multiple measurements of the same test in the same window, the value measured nearest to the target day will be selected to the window; if they are at the same distance to the target day, the latest one will be used. 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 5. Table for Visit Mapping for by Time Point Analysis

| Visit | Target Day | Analysis Time Window for Safety Data |
|--|------------|--------------------------------------|
| Screening (1) | <1 | <1 |
| Baseline (2) | 1 | 1 |
| Week 4 (3) | 29 | [2, 56] |
| Week 12 (4) | 85 | [57, 98] |
| Week 16 (5) | 113 | [99, 140] |
| Week 24 (6) | 169 | [141, 182] |
| Week 28 (7) | 197 | [183, 238] |
| Week 40 (8) | 281 | [239, 308] |
| Week 48 (9) | 337 | ≥309 |
| Target Day = Visit date – Drug start date + 1. | | |

For Placebo-controlled Phase, target day and relative window are relative to the first dose of IP.

The Apremilast-exposure Period encompasses all apremilast-exposure data, irrespective of when the apremilast exposure starts (at Week 0, 16, or 24). The scheduled study weeks for subjects who are originally randomized to the placebo group and transitioned to apremilast at Week 16/Visit 5 (due to early escape) or Week 24/Visit 6 will be adjusted to reflect the study weeks relative to the transition (i.e., study weeks on apremilast), and mapped to the (unadjusted) scheduled study weeks. Adjusted study weeks on apremilast will align with scheduled study weeks specified in the protocol.

This simple mapping will be ordered first as “minus 16” or “minus 24” from the scheduled visit depending on when the subject started apremilast, and then re-aligned to the

protocol scheduled visit. Scheduled weeks are: 0, 4, 12, 16, 24, 28, 40, 48; Placebo subjects who early escaped: 0, 8, 12, 24, 32; placebo subjects who transitioned at week 24: 0, 4, 16, 24. The records related to Week 8 are realigned to Week 4, 32 to Week 28.

Appendix B. Marked Abnormality Criteria

Table 6. Marked Abnormalities Criteria

| Category | Analyte (Unit) | Criteria |
|------------|--------------------------------------|---|
| Chemistry | Alanine Aminotransferase (U/L) | > 3×ULN |
| | Albumin (g/L) | < 30 |
| | Alkaline Phosphatase (U/L) | > 400 |
| | ALT/AST (U/L) | ALT or AST > 3×ULN |
| | Aspartate Aminotransferase (U/L) | > 3×ULN |
| | Bilirubin (µmol/L) | > 2×ULN |
| | Bilirubin (umol/L) and ALT/AST (U/L) | Bilirubin > 2×ULN and (ALT or AST > 3×ULN) |
| | Blood Urea Nitrogen (mmol/L) | > 15 |
| | Calcium (mmol/L) | 1. <1.8 2. >3.0 |
| | Cholesterol (mmol/L) | > 7.8 |
| | Creatinine (µmol/L) | > 1.7×ULN |
| | Glucose (mmol/L) | 1. < 2.8 2. > 13.9 |
| | Lactate Dehydrogenase (U/L) | > 3×ULN |
| | Magnesium (mmol/L) | > 1.2 |
| | Phosphate (mmol/L) | 1. < 0.64 2. > 1.60 |
| | Potassium (mmol/L) | 1. < 3.0 2. > 5.5 |
| | Sodium (mmol/L) | 1. < 130 2. > 150 |
| | Triglycerides (mmol/L) | > 3.4 |
| | Urate (umol/L) | Male: > 590; Female: > 480 |
| Hematology | Hemoglobin (g/L) | 1. Male: < 105; Female: < 85 2. Male: > 185; Female: > 170 |

| Category | Analyte (Unit) | Criteria |
|----------|----------------------------------|--|
| | Basophils (10 ⁹ /L) | > 3 |
| | Eosinophils (10 ⁹ /L) | 1. > 0.8 2. Increase ≥ 100% and value > 0.8 |
| | Hematocrit (Ratio) | 1. < 0.27 2. Male: > 0.55; Female: > 0.50 |
| | Leukocytes (10 ⁹ /L) | < 1.5 |
| | Lymphocytes (10 ⁹ /L) | < 0.8 |
| | Neutrophils (10 ⁹ /L) | < 1 |
| | Platelets (10 ⁹ /L) | 1. < 75 2. > 600 |
| | Hemoglobin A1C (%) | > 9 |