I9N-MC-FCAB Statistical Analysis Plan Version 1

A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 2 Study to Evaluate the Efficacy and Safety of LY3375880 in Adult Subjects with Moderate-to-Severe Atopic Dermatitis: The ADmIRe Study

NCT03831191

Approval Date: 27Nov2019

Statistical Analysis Plan: I9N-MC-FCAB(b): A Multicenter, Randomized, Double-Blind, Placebo-Controlled, Phase 2 Study to Evaluate the Efficacy and Safety of LY3375880 in Adult Patients with Moderate-to-Severe Atopic Dermatitis

2018-002401-56

Confidential Information

LY3375880

Study I9N-MC-FCAB (NCT03334396) is a Phase 2, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient, 60-week study designed to evaluate the efficacy and safety of LY3375880 50-mg, 150-mg, 300-mg, and 600-mg in adult patients with moderate to severe atopic dermatitis.

Eli Lilly and Company
Indianapolis, Indiana USA 46285
Protocol I9N-MC-FCAB(b)
Phase 2

Statistical Analysis Plan electronically signed and approved by Lilly on date provided below.

1. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

2. Revision History

Statistical Analysis Plan (SAP) Version 1 is based on Protocol I9N-MC-FCAB(b) and was approved prior to the first production transfer for the internal assessment committee.

3. Study Objectives

3.1. Primary Objective

The primary objective of this study is to test the hypothesis that LY3375880 is superior to placebo in the treatment of patients with moderate to severe atopic dermatitis (AD), as assessed by the proportion of patients achieving the validated Investigator's Global Assessment for AD (vIGA-AD, referred to throughout the SAP as IGA) of 0 or 1 with a ≥2-point improvement from baseline at Week 16.

3.2. Secondary Objectives

The secondary objectives of the study are the following:

Objectives	Endpoints
To compare the efficacy of LY3375880 to placebo as measured by improvement in signs and symptoms at Week 16 in the treatment of subjects with moderate- to- severe AD.	• Proportion of subjects achieving at Week 16: o EASI-50 o EASI-75 o EASI-90 o SCORAD-75 o SCORAD-90 o IGA of 0 • Mean change from baseline to Week 16 in: o EASI score o SCORAD • Proportion of subjects achieving IGA of 0 or 1 at Week 52
To characterize the pharmacokinetics of LY3375880 in subjects with moderate-to-severe atopic dermatitisAD	• Plasma pharmacokinetic data

3.3. Exploratory Objectives

The exploratory objectives of this study are the following:

Objectives/Endpoints
• To evaluate signs and symptoms, health outcome measures, and QoL measures (total scores, item scores, and derivations) at each time point collected. • To explore relationships between LY3375880 exposure and study endpoints. • To characterize post-induction loss of response and maintenance of response. • To evaluate PD effects of LY3375880 in skin biopsies and peripheral blood.

4. Study Design

4.1. Summary of Study Design

Study I9N-MC-FCAB(b) (FCAB) is a Phase 2, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study evaluating the efficacy and safety of LY3375880 subcutaneous (SC) 600 mg every 4 weeks (Q4W), 150 mg Q4W, and 50 mg Q4W as compared to placebo Q4W in adult subjects with moderate-to-severe AD. In addition, LY3375880 300mg Q4W will be evaluated during the maintenance period. The study duration will be up to 65 weeks over 4 study periods:

- Period 1: Screening Period lasting from 8 to 35 days prior to Week 0 (baseline, Visit 2).
- Period 2: Induction Period, lasting from Week 0 (baseline, Visit 2) through Week 16 (Visit 8), inclusive.
    - An internal assessment committee will conduct an unblinded interim analysis of efficacy safety data when approximately 40%-60% of patients complete Week 16 (Visit 8). This interim analysis will be conducted for internal decision making to trigger planning activities for future studies associated with LY3375880.
- Period 3: Maintenance Period, lasting from Week 20 (Visit 9) through Week 52 (Visit 17), inclusive. At Week 20, the following will occur:
    - Responders (defined as having achieved a 50% reduction in the Eczema Area and Severity Index score [EASI-50] response, regardless of whether rescue therapy has been initiated) will undergo the following re-randomizations:
        - Subjects on any dose of LY3375880 will be re-randomized at a 2:1 ratio to either maintain their current regimen or to receive placebo until loss of response (not achieving a 25% reduction in the EASI score [EASI-25] at a scheduled visit), at which point they will restart their induction period regimen.
        - Subjects on placebo will continue receiving placebo until loss of response, at which point they will receive LY3375880 300 mg Q4W.
    - Nonresponders will undergo the following re-randomizations/reassignments:
        - Subjects on 50 mg Q4W will be reassigned to 150 mg Q4W
        - Subjects on 150 mg Q4W will be reassigned to 600 mg Q4W
        - Subjects on 600 mg Q4W will remain on their current regimen
        - Subjects on placebo will receive LY3375880 300 mg Q4W.
    - Subjects with persistent disease activity defined as not achieving EASI-25 for 3 consecutive maintenance period visits, or 8 maintenance period weeks (whichever is longer) not already on 600 mg Q4W will be reassigned to 600 mg Q4W for the remainder of the study.
- Period 4: Post-Treatment Follow-Up Period, spanning approximately 8 weeks from Week 52 (Visit 17) to Week 60 (Visit 801).

Approximately 200 subjects will be randomized 1:1:1:1 at Week 0 to 1 of 4 treatment groups: placebo SC (Q4W), LY3375880 SC 600 mg (Q4W), LY3375880 150 mg (Q4W), or LY3375880

50 mg (Q4W). The study duration will be up to 65 weeks (Screening Period: Up to 5 weeks; Induction Period: 16 weeks; Maintenance Period: 36 weeks; Follow-up Period: 8 weeks).

Subjects will be stratified at randomization according to disease severity (IGA [3 versus 4]) and geographic region (Japan versus Rest of World).

Figure FCAB.4.1 illustrates the study design. The blinding procedure is described in the Protocol.

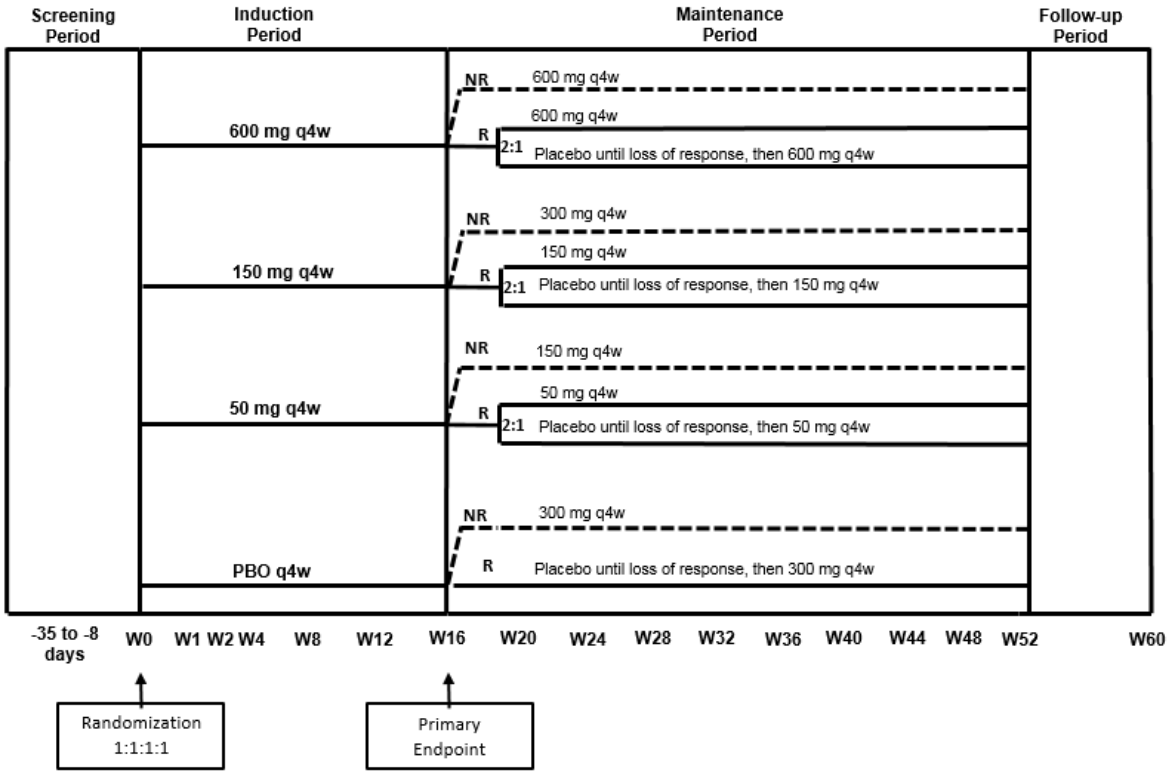

Figure FCAB.4.1. Illustration of study design for Clinical Protocol I9N-MC-FCAB.

Abbreviations: PBO = placebo; QD = once every four weeks; W = week.

4.2. Method of Assignment to Treatment

Subjects who meet all criteria for enrollment will be randomized 1:1:1:1 to double-blind treatment at Visit 2. Assignment to treatment groups will be determined by a computer-generated random sequence using an interactive web-response system (IWRS). The IWRS will be used to assign vials containing double-blind IP to each patient. Site personnel will confirm that they have located the correct vials by entering a confirmation number found on the vials into the IWRS. The IWRS will be used to assign IP to each patient. Site personnel will confirm that they have located the correct packages by entering a confirmation number found on packages into the IWRS.

To achieve between-group comparability, the randomization will be stratified by IGA (3 versus 4) and by geographic region (Japan versus Rest of World).

5. A Priori Statistical Methods

5.1. Determination of Sample Size

Study I9N-MC-FCAB(b) will aim to enroll approximately 200 patients ≥18 years of age. Assuming 36% and 8% of subjects achieve an IGA score of 0 or 1 (clear or almost clear skin) and ≥2-point improvement from baseline at Week 16 for LY3375880 and placebo, respectively, pairwise comparisons to placebo have at least 90% power using a 2-sided Fisher's exact test at the 0.05 significance level, with no adjustment for multiple comparisons.

Sample size and power estimates were obtained from R version 3.4.3.

5.2. General Considerations

This plan describes *a priori* statistical analyses for efficacy, health outcomes, and safety that will be performed.

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (Lilly). The statistical analyses will be performed using SAS® Version 9.4 or higher.

Not all displays described in this SAP will necessarily be included in the clinical study report (CSR). Not all displays will necessarily be created as a "static" display. Some may be incorporated into interactive display tools instead of or in addition to a static display. Any display described in this SAP and not included in the CSR would be available upon request.

For continuous measures, summary statistics will include sample size, mean, standard deviation (SD), minimum, median, and maximum. The 1st quartile and 3rd quartile may also be presented. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the database. The mean, 1st quartile, median, 3rd quartile, and SD will generally be reported to 1 more decimal place than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be 4 for any summary statistic. For categorical measures, summary statistics will include the sample size, frequency count, and percentage. Percentages will be presented to 1 decimal place unless otherwise stated. Percentages will not be presented for zero counts. Incidence rates and 95% confidence interval (CI) will be displayed for select safety endpoints.

Statistical tests of treatment effects and CIs will be performed at a 2-sided significance level of 0.05, unless otherwise stated.

All p-values will be rounded up to 3 decimal places. For example, any p-value strictly greater than 0.049 and less than or equal to 0.05 will be displayed as 0.050. This guarantees that on any printed statistical output, the unrounded p-value will always be less than or equal to the displayed p-value. A displayed p-value of 0.001 will always be understood to mean ≤0.001. Likewise, any p-value displayed as 1.000 will be understood to mean >0.999 and ≤1.

Data collected at early termination visits will be mapped to the next scheduled visit number for that patient if it falls within the visit window as discussed in Section 5.2.2. For by-visit summaries, only visits in which a measure was scheduled to be collected will be summarized. Any unscheduled visit data will be included at the patient-level listings. However, the data will

still be used in other analyses, including shift analyses for safety analytes, change from baseline to endpoint using modified last observation carried forward (mLOCF) analyses, and other categorical analyses.

5.2.1. Analysis Populations

Enrolled population: The enrolled population is defined as all participants who sign informed consent.

Intent-to-treat (ITT) population: The ITT population analysis set is defined as all randomized patients, even if the patient does not take the assigned treatment, does not receive the correct treatment, or otherwise does not follow the protocol.

Per-Protocol (PPS) population: The PPS populations is defined as all randomized patients who do not commit an Important Protocol Deviation (IPD) that could potentially compromise efficacy results.

Safety population: The safety population is defined as all randomized patients who receive at least 1 dose of their assigned double-blind study treatment..

Maintenance Period Population: All ITT patients who received at least one dose of study treatment and have entered the maintenance period at Week 16 (Visit 8).

Unless otherwise specified, the efficacy and health outcome analyses will be conducted on the ITT population (Gillings and Koch 1991), which seeks to preserve the benefits of randomization and avoid the issue of selection bias. Patients will be analyzed according to the treatment to which they were randomized.

Safety analyses will be done on the safety population. Patients will be analyzed according to the dosing regimen to which they were assigned. Analyses of the safety endpoints, many of which are incidence based, will include all patients in the safety population, unless specifically stated otherwise.

In the rare situation where a patient is Lost to Follow-up at the first postbaseline visit, but some safety data exists (eg, unscheduled laboratory assessments) after first dose of study drug, a listing of the data or a patient profile will be provided.

5.2.2. Definition of Baseline and Postbaseline Measures

The baseline value for the efficacy and health outcomes for the Induction and Maintenance Periods is defined as the last non-missing measurement on or prior to the date of first study drug administration (expected at Week 0, Visit 2). The baseline value for the daily diary assessments (Itch NRS, Skin Pain NRS) is defined as the average of the non-missing assessments in the last 7 days prior to the date of first study drug administration (expected at Week 0, Visit 2). Criteria for derivation of the weekly average requires that there must be at least 4 non-missing measurements in the 7 days indicated, otherwise, an expanded window of 14 days prior to first study drug administration, if available, may be utilized in order to obtain the most recent 4 non-missing measurements prior to first study drug administration (ie, Day -14 to Day -1). If there are not at least 4 non-missing assessments collected prior to the date of first study drug

administration using the aforementioned method, then the baseline will be designated as missing. Other definitions of baseline value, such as the last measurement on or prior to rescue, may be used to conduct additional supporting analyses.

Postbaseline measurements are collected after study drug administration through Week 60 (Visit 801) or early discontinuation visit. For data collected in the electronic Clinical Outcomes Assessment (eCOA) tablet (including Patient-Reported Outcomes [PRO] and Clinician-Reported Outcomes [ClinRO]) and related to efficacy assessments, unscheduled postbaseline visits that fall within the visit windows defined by Lilly will be summarized in the by-visit analyses if there is no scheduled visit available. That is, a ±2 day window is applied to Visit 3 (Week 1), Visit 4 (Week 2), and Visit 5 (Week 4), and a ±4 day window is applied to Visit 6 (Week 8), Visit 7 (Week 12), and Visit 8 (Week 16), as well as to every visit in the Maintenance Period. If there is more than 1 unscheduled visit within the defined visit window and no scheduled visit is available, the unscheduled visit closest to the scheduled visit date will be used.

Postbaseline value of daily diary assessments collected after study drug administration through Week 60 (Visit 801) or early discontinuation visit will be defined as the average of the last 7 days prior to the date of the study visit and on or after the date of the previous visit. If there are less than 4 non-missing assessments in the 7-day window, the window will be extended 1 day at a time, up to 10 days, until 4 non-missing assessments are found, but the window will not be extended beyond the date of the previous visit. If at least 4 non-missing measurements are still not found, the visit will be missing. If a 7-day window is not present between visits, the postbaseline value will be the average of at least 4 non-missing assessments prior to the current visit, and on or after the previous visit. If at least 4 non-missing assessments are not available and the visit window is 5 or 6 days in length, then at least 3 non-missing assessments prior to the current visit, and on or after the previous visit will be acceptable.

5.2.2.1. Induction Period Considerations

The baseline value for the safety analyses for the Induction Period is defined as the last non-missing measurement on or prior to the date of first study drug administration (expected at Week 0, Visit 2). Other definitions of baseline value, such as the last measurement on or prior to rescue, may be used to conduct additional supporting analyses.

5.2.2.2. Maintenance Period Considerations

The baseline value for the safety analyses will be the last available value before the intial dose in the Maintenance period (expected at Week 16, Visit 8). Other definitions of baseline value, such as the last measurement on or prior to reassignment to active treatment, may be used to conduct additional supporting analyses.

5.2.3. Analysis Methods

The main analysis of categorical efficacy variables and health outcomes variables will use a logistic regression analysis with region, baseline disease severity (IGA) and treatment group in the model. The p-value and 100(1-alpha)% CI for the odds ratio from the logistic regression model are used for primary statistical inference. The difference in percentages and 100(1-alpha)% CI of the difference in percentages using the Newcombe-Wilson without continuity

correction will be reported. The p-value from the Fisher's exact test will also be produced as a secondary analysis.

The main analysis for all continuous efficacy and health outcomes variables will use mixed model repeated measures (MMRM) analysis. The MMRM model will use a restricted maximum likelihood (REML) estimation. The model will include treatment, region, baseline disease severity, visit, and treatment-by-visit-interaction as fixed categorical effects and baseline and baseline-by-visit-interaction as fixed continuous effects. An unstructured (co)variance structure will be used to model the between- and within-patient errors. If this analysis fails to converge, the heterogeneous autoregressive [ARH(1)], followed by the heterogeneous compound symmetry (CSH), followed by the heterogeneous Toeplitz (TOEPH) will be used. The Kenward-Roger method will be used to estimate the degrees of freedom. Type III tests for the least-squares means (LSM) will be used for the statistical comparison. The LSM difference, standard error, p-value and 100(1-alpha)% CI will also be reported.

Fisher's exact test will be used for the adverse events (AEs), discontinuation, and other categorical safety data for between-treatment group comparisons. Continuous vital signs, body weight, and other continuous safety variables, including laboratory variables will be analyzed by an ANCOVA with treatment group and baseline value in the model. The significance of within-treatment group changes from baseline will be evaluated by testing whether or not the treatment group LSM changes from baseline are different from zero; the standard error for the LSM change will also be displayed. Differences in LSM will be displayed, with the p-value associated with the LSM comparison to placebo and a 95% CI on the LSM difference also provided. In addition to the LSMs for each group, the within-group p-value for the change from baseline will be displayed. Treatment-emergent high/low for categorical laboratory safety analyses will also be produced.

5.3. Covariate Adjustment

The randomization to treatment groups at Week 0 (Visit 2) is stratified by disease severity (IGA) and geographic region as described in Section 4.1. Unless otherwise specified, the statistical analysis models will adjust for disease severity and geographical region. When an MMRM analysis is performed, baseline value and baseline-by-visit interactions will be included as covariates.

5.4. Handling of Dropouts or Missing Data

Nonresponder imputation (for categorical variables) and MMRM (for continuous variables) will be the primary methods used to handle missing data. Subjects who receive rescue therapy will be considered missing at rescue and all subsequent visits.

5.4.1. Nonresponder Imputation

For the analysis of categorical efficacy and health outcomes variables such as IGA(0,1) and EASI50/75/90, the primary method to impute missing data will be a nonresponder imputation (NRI) method.

All categorical endpoints will utilize the NRI method to patients who permanently discontinued study drug or were rescued.

5.4.2. Mixed Model for Repeated Measures

For the continuous secondary and exploratory efficacy and health outcome variables, such as EASI score and SCORAD score, MMRM analyses will be performed to mitigate the impact of missing data. This approach assumes that missing observations are missing-at-random (missingness is related to observed data) during the study and takes into account both the missingness of data and the correlation of the repeated measurements.

All continuous endpoints will utilize MMRM to patients who permanently discontinued study drug or who were rescued.

5.5. Multicenter Studies

This study will be conducted by multiple investigators at multiple sites internationally. The countries will be categorized into geographic regions, as described in Section 4.2.

For the analysis of the primary endpoint, treatment-by-region interaction will be added to the logistic regression model as a sensitivity analysis and results from this model will be compared to the primary model (without the interaction effect). If the treatment-by-region interaction is significant at a 2-sided α level of 0.1, the nature of this interaction will be inspected as to whether it is quantitative (ie, the treatment effect is consistent in direction across all regions but not in size of treatment effect) or qualitative (the treatment is beneficial in some but not all regions). If the treatment-by-region interaction effect is found to be quantitative, results from the primary model will be presented. If the treatment-by-region interaction effect is found to be qualitative, further inspection will be used to identify in which regions LY3375880 is found to be more beneficial.

5.6. Multiple Comparisons/Multiplicity

No adjustments for multiplicity will be made for any analyses in this study.

5.7. Patient Disposition

An overview of patient populations will be summarized by treatment group. Frequency counts and percentages of patients excluded prior to randomization by primary reason for exclusion will be provided for patients who failed to meet study entry requirements during screening.

Patient disposition through Week 16 will be summarized using the ITT population. Frequency counts and percentages of patients who complete the Induction Period and Maintenance Period or discontinue early from the study will be summarized separately by treatment group for patients who are not rescued and for patients who are rescued, along with their reason for study discontinuation. Frequency counts and percentages of patients who complete treatment or discontinue treatment early will also be summarized separately by treatment group for patients who are not rescued and for patients who are rescued, along with their reason for treatment discontinuation.

A listing of patient disposition will be provided for all randomized patients, with the extent of their participation in the study and the reason for discontinuation. A listing of all randomized patients with their treatment assignment will also be provided.

5.8. Patient Characteristics

Patient characteristics including demographics and baseline characteristics will be summarized descriptively by treatment group for the ITT population. Historical illnesses and pre-existing conditions will be summarized descriptively by treatment group for the ITT population. Descriptive statistics including number of patients (n), mean, SD, minimum, median, and maximum will be provided for continuous measures, and frequency counts and percentages will be tabulated for categorical measures. No formal statistical comparisons will be made among treatment groups unless otherwise stated.

5.8.1. Demographics

Patient demographics will be summarized as described above. The following demographic information will be included:

- Age
- Age group (<65 vs. ≥65)
- Age group (<65, ≥65 to <75, ≥75 to <85, ≥85)
- Gender (male, female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
- Region (Japan, Rest of World)
- Country
- Weight (kg)
- Height (cm)
- Body mass index (kg/m^2)

A listing of patient demographics will also be provided for the ITT population.

5.8.2. Baseline Disease Characteristics

The following baseline disease information (but not limited to only these) will be categorized and presented for baseline AD clinical characteristics, baseline health outcome measures, and other baseline demographic and disease characteristics as described above:

- Duration since AD diagnosis (years)
- Duration since AD diagnosis category (0 to <2 years, 2 to <5 years, 5 to <10 years, 10 to <20 years, ≥20 years)
- Validated Investigator's Global Assessment for AD (IGA) score
- Eczema Area and Severity Index (EASI) score
- SCORing Atopic Dermatitis (SCORAD)
- Body Surface Area (BSA) affected by AD
- Patient-Oriented Eczema Measure (POEM)
- Itch Numerical Rating Scale (NRS)

- Dermatology Life Quality Index (DLQI)
- Skin Pain NRS
- Immunoglobulin E (IgE): intrinsic(<200 kU/I) or extrinsic (≥200 kU/I)

5.8.3. Historical Illness and Pre-existing Conditions

Pre-existing Conditions and Medical History will be coded by Medical Dictionary for Regulatory Activities (MedDRA, most current available version). Data will be listed and summarized by treatment, system organ class (SOC), and preferred term (PT).

Pre-existing condition is defined as a condition with a start date on or before the informed consent date. Pre-existing conditions will be used to derive treatment-emergent adverse events (TEAEs).

5.9. Treatment Compliance

In the Induction Phase, patient compliance with study medication will be assessed from Week 0 (Visit 2) to Week 16 (Visit 8) or Early Termination using the ITT population. In the Maintenance Phase, patient compliance with study medication will be assessed from Week 16 (Visit 8) to Week 52 (visit 17) or Early Termination using the Maintenance Period population.

All doses of study medication will be administered at the study site by study personnel. Patients who miss 20% of the expected doses through Week 16(Visit 8, exclusive) or through Week 52 (Visit 17) will be excluded from the PPS population in the Induction Period and Maintenance Period, respectively.

Descriptive statistics for percent compliance and non-compliance rate will be summarized for the ITT population by treatment group for Week 0 through Week 16, and for the Maintenance Period population by treatment group for Week 16 through Week 52, with data up to rescue. .Percent compliance will be listed by patient for Week 0 through Week 16 and for Week 16 through Week 52, with data up to rescue.

5.9.1. Rescue Treatment

Summaries of patients receiving rescue therapy at each visit will be provided.

5.10. Previous and Concomitant Therapy

Summaries of previous and concomitant medications will be based on the ITT population.

Concomitant therapy for the treatment period is defined as therapy that starts before or during the treatment period and ends during the treatment period or is ongoing (has no end date or ends after the treatment period). Should there be insufficient data to make this comparison (for example, the concomitant therapy stop year is the same as the treatment start year, but the concomitant therapy stop month and day are missing), the medication will be considered as concomitant for the treatment period.

5.11. Efficacy and Health Outcome Analyses

The general methods used to summarize efficacy data, including the definition of baseline value for assessments are described in Section 5.1.

Table FCAB.5.1 includes the descriptions and derivations of the primary, secondary, and exploratory efficacy and health outcomes.

Table FCAB.5.2 provides the detailed analyses including analysis type, method and imputation, population, time point, and comparisons for efficacy and health outcome analyses.

Table FCAB.5.1. Description and Derivation of Primary, Secondary and Exploratory Efficacy Outcomes

Measure	Description	Variable	Derivation / Comment	Imputation Approach if with Missing Components
Validated Investigator's Global Assessment for AD (IGA)	The validated Investigator's global assessment of the patient's overall severity of their AD, based on a static, numeric 5-point scale from 0 (clear) to 4 (severe). The score is based on an overall assessment of the degree of erythema, papulation/induration, oozing/crusting, and lichenification.	▪ IGA score	Single item. Range: 0 to 4 0 represents "clear" 4 represents "severe"	Single item, missing if missing.
		▪ Change from baseline in IGA score	Change from baseline: observed IGA score – baseline IGA score	Missing if baseline or observed value is missing.
		▪ IGA [0,1] with ≥2-point improvement ▪ IGA [0]	▪ Observed score of 0 or 1 and change from baseline ≤-2 ▪ Observed score of 0	▪ Missing if baseline or observed value is missing. ▪ Single item, missing if missing.
Eczema Area and Severity Index (EASI)	The EASI assesses objective physician estimates of 2 dimensions of atopic dermatitis – disease extent and clinical signs (Hanifin et al 2001) – by scoring the extent of disease (percentage of skin affected: 0 = 0%; 1 = 1-9%; 2 = 10-29%; 3 = 30-49%; 4 = 50-69%; 5 = 70-89%; 6 = 90-100%) and the severity of 4 clinical signs (erythema, edema/papulation, excoriation, and lichenification) each on a scale of 0 to 3 (0 = none, absent; 1 = mild; 2 = moderate; 3 = severe) at 4 body sites (head and neck, trunk, upper limbs, and lower limbs). Half scores are allowed. Each body site will have a score that ranges from 0 to 72, and the final EASI score will be obtained by weight-averaging these 4 scores. Hence, the final EASI score will range from 0 to 72 for each time point.	▪ EASI score	Derive EASI region score for each of head and neck, trunk, upper limbs, and lower limbs as follows: $EASI_{region}$ = (Erythema + edema/papulation + Excoriation + Lichenification) *(value from percentage involvement), where erythema, edema/papulation, excoriation, and lichenification are evaluated on a scale of 0 to 3 and value from percentage involvement is on a scale of 0 to 6. Then total EASI score is as follows: $EASI = 0.1*EASI_{head\ and\ neck} + 0.3*EASI_{trunk} + 0.2*EASI_{upper\ limbs} + 0.4*EASI_{lower\ limbs}$	N/A – partial assessments cannot be saved.
		▪ Change from baseline in EASI score ▪ Percent change from baseline EASI score	Change from baseline: observed EASI score – baseline EASI score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.
		▪ EASI75	% Improvement in baseline ≥75% or % change from baseline ≤-75%	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if with Missing Components
		▪ EASI90	% Improvement in baseline ≥90% or % change from baseline ≤-90%	Missing if baseline or observed value is missing.
Body Surface Area (BSA) Affected by AD	Body surface area affected by AD will be assessed for 4 separate body regions and is collected as part of the EASI assessment: head and neck, trunk (including genital region), upper extremities, and lower extremities (including the buttocks). Each body region will be assessed for disease extent ranging from 0% to 100% involvement. The overall total percentage will be reported based off of all 4 body regions combined, after applying specific multipliers to the different body regions to account for the percent of the total BSA represented by each of the 4 regions.	▪ BSA score	Use the percentage of skin affected for each region (0 to 100%) in EASI as follows: $BSA = BSA_{head\ and\ neck}/100/0.10 + BSA_{trunk}/100/0.0333 + BSA_{upper\ limbs}/100/0.05 + BSA_{lower\ limbs}/100/0.025$	N/A – partial assessments cannot be saved.
		▪ Change from baseline in BSA score	Change from baseline: observed BSA score – baseline BSA score	Missing if baseline or observed value is missing.
SCORing Atopic Dermatitis (SCORAD)		▪ SCORAD score	SCORAD = A/5 + 7B/2 + C, where A is extent of disease, range 0-100 B is disease severity, range 0-18 C is subjective symptoms, range 0-20	Missing if components A and B are missing or if component C is missing. Partial assessments performed by physician cannot be saved and partial assessments performed by patient cannot be saved.
		▪ Change from baseline in SCORAD score ▪ Percent change from baseline in SCORAD score	Change from baseline: observed SCORAD score – baseline SCORAD score % change from baseline: $100 \times \frac{Observed\ score - Baseline}{Baseline}$	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if with Missing Components
	The SCORing Atopic Dermatitis (SCORAD) index uses the rule of nines to assess disease extent (head and neck 9%; upper limbs 9% each; lower limbs 18% each; anterior trunk 18%; back 18%; and genitals 1%). It evaluates 6 clinical characteristics to determine disease severity: (1) erythema, (2) edema/papulation, (3) oozing/crusts, (4) excoriation, (5) lichenification, and (6) dryness on a scale of 0 to 3 (0=absence, 1=mild, 2=moderate, 3=severe). The SCORAD index also assesses subjective symptoms of pruritus and sleep loss in the last 72 hours on visual analogue scales (VAS) of 0 to 10 where 0 is no itch or sleep loss and 10 is worst imaginable itch or sleep loss. These 3 aspects: extent of disease, disease severity, and subjective symptoms combine to give a maximum possible score of 103 (Stalder et al. 1993; Kunz et al. 1997; Schram et al. 2012).	▪ SCORAD75	% Improvement in baseline ≥75%: % change from baseline ≤-75	Missing if baseline or observed value is missing.
		▪ SCORAD90	% Improvement in baseline ≥90%: % change from baseline ≤-90	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if with Missing Components
Itch Numeric Rating Scale (NRS)	The Itch Numeric Rating Scale (NRS) is a patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no itch" and 10 representing "worst itch imaginable." Overall severity of a patient's itching is indicated by selecting the number that best describes the worst level of itching in the past 24 hours (Naegeli et al. 2015; Kimball et al. 2016). Refer to Section 5.2.2 for details on how to calculate the weekly score which will be used in the continuous analysis.	▪ Itch NRS score	Single item; range 0-10. Refer to Section 5.2.2 on how to derive the visit score.	Refer to Section 5.2.2 on how to derive the visit score.
		▪ Change from baseline in Itch NRS	Change from baseline: observed Itch NRS score – baseline Itch NRS score	Missing if baseline or observed value is missing. Refer to Section 5.2.2 on how to derive the visit score.
Skin Pain Numeric Rating Scale (NRS)	Skin Pain NRS is a patient-administered, 11-point horizontal scale anchored at 0 and 10, with 0 representing "no pain" and 10 representing "worst pain imaginable." Overall severity of a patient's skin pain is indicated by selecting the number that best describes the worst level of skin pain in the past 24 hours Refer to Section 5.2.2 for details on how to calculate the weekly score which will be used in the continuous analysis.	▪ Skin Pain NRS score	Single item; range 0 to 10. Refer to Section 5.2.2 on how to derive the visit score.	Refer to Section 5.2.2 on how to derive the visit score.
		▪ Change from baseline in Skin Pain NRS	Change from baseline: observed skin pain score – baseline skin pain score	Missing if baseline or observed value is missing.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if with Missing Components
		▪ Frequency of No Pain (Skin pain NRS = 0)	Count of observed value = 0 starting on the day of the previous study visit to the day prior to the subsequent study visit. This only applies to scheduled study visits and thus will be calculated for the following visit intervals: Week 0 to Week 1, Week 1 to Week 2, Week 2 to Week 4, Week 4 to Week 8, Week 8 to Week 12 and Week 12 to Week 16. The number of observations with value = 0 will then be pro-rated for the interval; intervals of 1 week will adjusted to 7 days, 2 weeks to 14 days and 4 weeks to 28 days. For example, if the patient has 5 assessments with a value of 0 from Week 12 to Week 16 with 30 of the 35 days occurring between the visits having non-missing values, the pro-rated count will be 5/30*28=4.7.	If patients do not have at least an average of 4 values for each week in the visit interval (eg, at least 4 non-missing assessments for Week 1 to Week 2, at least 8 non-missing assessments from Week 2 to Week 4, at least 16 non-missing assessments from Week 12 to 16), the score is missing.
Patient- Oriented Eczema Measure (POEM)		▪ POEM score	POEM total score: sum of questions 1 to 7, Range 0 to 28.	If a single question is left unanswered, then that question is scored as 0. If more than one question is unanswered, then the tool is not scored. If more than one response is selected, then the response with the highest score is used.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if with Missing Components
	The POEM is a simple, 7-item, patient-administered scale that assesses disease severity in children and adults. Patients respond to questions about the frequency of 7 symptoms (itching, sleep disturbance, bleeding, weeping/oozing, cracking, flaking, and dryness/roughness) over the last week. Response categories include "No days," "1-2 days," "3-4 days," "5-6 days," and "Every day" with corresponding scores of 0, 1, 2, 3, and 4, respectively. Scores range from 0-28 with higher total scores indicating greater disease severity (Charman et al. 2004).	▪ Change from baseline in POEM score	Change from baseline: observed POEM score – baseline POEM score	Missing if baseline or observed value is missing.
Dermatology Life Quality Index (DLQI)		▪ Symptoms and feelings domain	Sum of questions 1 and 2, range 0 to 6.	N/A – partial assessments cannot be saved.
		▪ Daily activities domain	Sum of questions 3 and 4, range 0 to 6.	N/A – partial assessments cannot be saved.
		▪ Leisure domain	Sum of questions 5 and 6, range 0 to 6.	N/A – partial assessments cannot be saved.
		▪ Work and school domain	Sum of questions 7 and 7B (if it is answered), range 0 to 3. Responses of "yes" and "no" on Question 7 are given scores of 3 and 0 respectively. If Question 7 is answered "no" then Question 7b is answered with "a lot", "a little", "not at all" getting scores of 2, 1, 0 respectively.	N/A – partial assessments cannot be saved.
		▪ Personal relationships domain	Sum of questions 8 and 9, range 0 to 6.	N/A – partial assessments cannot be saved.
		▪ Treatment domain	Question 10, range 0 to 3.	N/A – partial assessments cannot be saved.

Measure	Description	Variable	Derivation / Comment	Imputation Approach if with Missing Components
	The Dermatology Life Quality Index (DLQI) is a simple, patient-administered, 10-item, validated, quality-of-life questionnaire that covers 6 domains including symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment. The recall period of this scale is over the "last week." Response categories include "a little," "a lot," and "very much," with corresponding scores of 1, 2, and 3, respectively, and "not at all," or unanswered ("not relevant") responses scored as 0. Scores range from 0-30 with higher scores indicating greater impairment of quality of life. A DLQI total score of 0 to 1 is considered as having no effect on a patient's health-related QoL (Hongbo et al. 2005), and a 4-point change from baseline is considered as the minimal clinically important difference threshold (Khilji et al. 2002; Basra et al. 2015).	▪ DLQI total score	DLQI total score: sum of all six DLQI domain scores, range 0 to 30.	N/A – partial assessments cannot be saved.
		▪ Change from baseline in DLQI	Change from baseline: observed DLQI score – baseline DLQI score	Missing if baseline or observed value is missing.

Table FCAB.5.2. Description of Primary, Secondary and Exploratory Efficacy Analyses

Measure	Variable	Analysis Method (Section 5.2.3)	Population (Section 5.2.1)	Comparison/Time Point	Analysis Type
Validated Investigator's Global Assessment for AD (IGA)	• Proportion of patients achieving IGA [0,1] with a ≥2-point improvement	Logistic regression using NRI	ITT; PPS	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Primary analysis
			Maintenance	LY3375880 600-mg or LY3375880 300-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 52	Secondary analysis
	• Proportion of patients achieving IGA [0]	Logistic regression using NRI	ITT	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Secondary analysis
Eczema Area and Severity Index (EASI)	• EASI score • Change from baseline in EASI score • Percent change from baseline in EASI score	MMRM	ITT; PPS	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Secondary analysis
	• Proportion of patients achieving EASI50 • Proportion of patients achieving EASI75 • Proportion of patients achieving EASI90	Logistic regression using NRI	ITT; PPS	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Secondary analysis
Body Surface Area (BSA) Affected by AD	• BSA score • Change from baseline in BSA score	MMRM	ITT	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Exploratory analysis
SCORing Atopic Dermatitis (SCORAD)	• SCORAD score • Change from baseline in SCORAD score • Percent change from baseline in SCORAD score	MMRM	ITT	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Secondary analysis
	• Proportion of patients achieving SCORAD75 • Proportion of patients achieving SCORAD90	Logistic regression using NRI	ITT; PPS	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Secondary analysis

Measure	Variable	Analysis Method (Section 5.2.3)	Population (Section 5.2.1)	Comparison/Time Point	Analysis Type
Itch NRS	• Itch NRS score • Change from baseline in Itch NRS score	MMRM	ITT; PPS	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Exploratory Analysis
	• Proportion of patients with at least a 4-point improvement in Itch NRS	Logistic regression using NRI	ITT; PPS (those with baseline Itch NRS of 4 or greater)	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Exploratory Analysis
Skin Pain Numeric Rating Scale (NRS)	• Skin Pain NRS score • Change from baseline in Skin Pain NRS score	MMRM	ITT; PPS	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Exploratory Analysis
Patient-Oriented Eczema Measure (POEM)	• POEM score • Change from baseline in POEM score	MMRM	ITT	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Exploratory Analysis
Dermatology Life Quality Index (DLQI)	• DLQI total score • Change from baseline in DLQI	MMRM	ITT	LY3375880 600-mg or LY3375880 150-mg or LY3375880 50-mg vs PBO; Week 16	Exploratory Analysis

Abbreviations: ITT = intent-to-treat; MMRM = mixed model repeated measures; NRI = nonresponder imputation; NRS = Numeric Rating Scale; PBO = placebo; PPS = per protocol set.

5.11.1. Primary Outcome and Methodology

The validated Investigator's Global Assessment for AD (IGA) uses the clinical characteristics of erythema, papulation/induration, oozing/crusting and lichenification to produce a single-item score ranging from 0 to 4. The primary analysis of the study is to test the hypotheses that LY3375880 600-mg, LY3375880 150-mg, or LY3375880 50-mg is superior to placebo when evaluating the proportion of patients achieving IGA of 0 or 1 with a ≥2-point improvement from baseline at Week 16 using the ITT population. A logistic regression analysis as described in Section 5.2.3 will be used for the comparisons. The odds ratio, the corresponding 95% CIs and p-value, as well as the treatment differences and the corresponding 95% CIs, will be reported. Missing data will be imputed using the NRI method described in Section 5.4.1.

5.11.2. Secondary and Exploratory Efficacy and Health Outcomes Analyses

The secondary and exploratory efficacy analyses, as well as the health outcomes analyses, are detailed in Table FCAB.5.1. There will be no adjustment for multiple comparisons forother analyses.

5.12. Bioanalytical and Pharmacokinetic/Pharmacodynamic Methods

Pharmacokinetic, Pharmacodynamic and Biomarker analyses to address secondary and exploratory objectives of this study will be described by Lilly in separate PK/PD and Biomarker analysis plans.

5.13. Safety Analyses

The general methods used to summarize safety data, including the definition of baseline value are described in Section 5.2.2.

Safety analyses will include data after rescue, unless otherwise stated, and patients will be analyzed according to the investigational product to which they were randomized at Week 0 (Visit 2). Additional analyses may be conducted using data after rescue to systemic therapy for some safety topics such as systemic treatment-emergent adverse events (TEAEs), and serious adverse events (SAEs). Safety analyses will take place using the safety population defined in Section 5.2.1.

Safety topics that will be addressed include the following: AEs, clinical laboratory evaluations, vital signs and physical characteristics, Columbia Suicide Severity Rating Scale (C-SSRS), the Self-Harm Supplement Form, safety in special groups and circumstances, including adverse events of special interest (AESI) (see Section 5.13.5), and investigational product interruptions.

Unless otherwise specified, by-visit summaries will include planned on-treatment visits. For tables that summarize events (such as AEs, categorical lab abnormalities, shift to maximum value), post-last dose follow-up data will be included. Follow-up data is defined as all data occurring up to 30 days (planned maximum follow-up time) after last dose of treatment including rescue, regardless of study period except for deaths and malignancies. For deaths and

malignancies, all available follow-up data up to the end of the study will be included. Listings will include all safety data.

For selected safety assessments, descriptive statistics may be presented for the last measure observed during post-treatment follow-up (up to 30 days after the last dose of treatment including rescue, regardless of study period).

5.13.1. Extent of Exposure

Duration of exposure (in days) to study drug will be summarized for the safety population in the Induction Period and for the maintenance population in the Maintenance Period by treatment group using descriptive statistics (n, mean, SD, minimum, 1st quartile, median, 3rd quartile, maximum).

For subjects without drug interruption, the duration of exposure for each patient is the date of last dose – date of first dose +1. If subjects have drug interruption, the duration of exposure for each patient is the (Treatment end date before drug interruption – date of first dose +1) + (date of last dose – treatment re-start date +1).

5.13.2. Adverse Events

Adverse events are recorded in the eCRFs. Each AE will be coded to system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA) version that is current at the time of database lock. Severity of AEs is recorded as mild, moderate, or severe.

A TEAE is defined as an event that either first occurred or worsened in severity after the first dose of study treatment and on or prior to the last visit date during the analysis period. The analysis period is defined as the treatment period plus 30 days off-drug follow-up time, except for deaths and malignancies for which the analysis period includes the treatment period and off-drug follow-up time.

Adverse events are classified based upon the MedDRA PT. The MedDRA Lowest Level Term (LLT) will be used in defining which events are treatment-emergent. The maximum severity for each LLT during the baseline period up to first dose of the study medication will be used as baseline. If an event with missing severity is preexisting during the baseline period, and persists during the treatment period, then the baseline severity will be considered mild for determining TEAE (that is, the event is treatment-emergent if the severity is coded moderate or severe postbaseline and not treatment-emergent if the severity is coded mild postbaseline). If an event occurring postbaseline has a missing severity rating, then the event is considered treatment-emergent. The day and time for events where onset is on the day of the first dose of study treatment will both be used to distinguish between pretreatment and posttreatment in order to derive treatment-emergence. Should there be insufficient data for AE start date to make this comparison (for example, the AE start year is the same as the treatment start year, but the AE start month and day are missing), the AE will be considered treatment-emergent.

In an overview table, the number and percentage of subjects who experienced a TEAE or serious adverse event (SAE), died due to an AE, or discontinued from study due to an AE will be summarized by treatment.

The frequency and percentages of patients with TEAEs will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC. A common TEAEs (at least reported by 5% of patients) summary will also be provided. As an additional table or figure, the percentages of patients with TEAEs will be summarized by treatment using MedDRA PT (without regard to SOC). Events will be ordered by decreasing frequency within PT. The frequency and percentages of patients with TEAEs by maximum severity will also be summarized by treatment using MedDRA PT nested within SOC.

Fisher's exact test will be performed to compare percentages for each LY3375880 dose group and Placebo at both the SOC and PT levels.

Listings will be presented for all AEs, all SAEs, all AEs leading to death, all AEs leading to study drug temporary interruption, and all AEs leading to discontinuation from the study.

5.13.2.1. Serious Adverse Event Analyses

The number and percentage of patients who experienced an SAE (including the SAEs that led to death) during the study drug period will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC.

5.13.2.2. Other Significant Adverse Events

The number and percentage of patients who permanently discontinued from study treatment due to an AE (including AEs that led to death) during the study drug period will be summarized by treatment using MedDRA PT nested within SOC. Events will be ordered by decreasing frequency within SOC.

Fisher's exact test will be performed to compare percentages for each LY3375880 dose group and Placebo at both the SOC and PT levels.

5.13.3. Clinical Laboratory Evaluation

Laboratory analyses will include planned analytes only. Planned analytes are those specified in the protocol. Summaries will be provided in both Système International (SI) and United States (US) conventional (CN) units (when different). Limits from the performing lab will be used to define low and high. Analyses of laboratory values will be produced using the measurements collected at scheduled time points. All unscheduled assessments will appear only in data listings.

- **Box plots for observed values:** Values at each visit (starting at randomization) will be displayed in box plots for patients who have both a baseline and a result for the specified visit. Unplanned measurements will be excluded. Original-scale data will be used for the display. These box plots will be used to evaluate trends over time and to assess a potential impact of outliers on central tendency summaries. The box plot will be a

notched box for each treatment with outliers displayed, individual measurements outside of reference limits will also be displayed using distinct symbols overlaying the box plot, and descriptive summary statistics will be included in a table below the box plot.

- **Box plots for change values:** Change from baseline to each visit will be displayed in box plots for patients who have both a baseline and a result for the specified visit. Change from baseline to last observation will also be summarized and analyzed for patients who have both baseline and at least 1 postbaseline result. Baseline will be the last non-missing observation in the baseline period. The last non-missing observation in the Dosing Period will be used as the last observation. Unplanned measurements will be excluded. The box plot will be a notched box for each treatment with outliers displayed, change from baseline to last observation will be summarized within the box plot of changes (rightmost column), and descriptive summary statistics will be included in a table below the box plot, along with a p-value using the ANCOVA model containing terms for treatment and the continuous covariate of baseline measurement. Type III sums of squares will be used. The significance of within-treatment group changes from baseline will be evaluated by testing whether the treatment group LSMean changes from baseline are different from zero using a t-statistic. In addition to the LSMeans and tests, the SD, minimum, Q1, median, Q3, and maximum will be displayed.
- **Outlier/shift displays focusing on low values (where low values are of interest):** A scatterplot, a shift table, and a shift to low table will be created for both the Induction and Maintenance Period. Unplanned measurements will be included. The scatterplot will plot the baseline value versus the minimum value during the study period. Lines indicating the reference limits will be included. In cases where limits vary across demographic characteristics, lines indicating the most common limit will be displayed. The shift table will include the number and percentage of patients within each baseline category (low, normal, high, or missing) versus each postbaseline category (minimum value is low, normal, high, or missing) by treatment. Patients in the safety population will be included in the shift table. The shift from normal/high to low table will include the number and percentage of patients by treatment whose baseline result is normal or high and whose minimum treatment result is low. Patients whose baseline result is normal or high and have at least 1 post-baseline result during the study period are included. The Fisher's exact test will be used to compare percentages of patients who shift from normal/high to low between treatments.
- **Outlier/shift displays focusing on high values (where high values are of interest):** The same approach described for low values will be used, except maximum values will replace minimum values.

Alanine aminotransferase (ALT), aspartate aminotransferase (AST), and total bilirubin will not be included in this analysis as different ranges will be used as described in Section 5.13.5.

5.13.4. Vital Signs and Other Physical Findings

Vital signs and physical characteristics include systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse, weight, and BMI. Original-scale data will be analyzed. When these

parameters are analyzed as continuous numerical variables, unplanned measurements will be excluded. When these parameters are analyzed as categorical outcomes and/or treatment-emergent abnormalities, planned and unplanned measurements will be included.

The planned analyses described for the laboratory analytes in Section 5.13.3 will be used to analyze the vital signs and physical characteristics, except for the inclusion of a threshold for change in addition to a limit for the definition of treatment-emergent.Table FCAB.5.3 defines the low and high baseline values as well as the criteria used to define treatment-emergence based on post-baseline values.

Table FCAB.5.3. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressure and Pulse Measurement, and Categorical Criteria for Weight Changes for Adults

Parameter	Low mmHg	High mmHg
Systolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤90 and decrease from baseline ≥20	≥140 and increase from baseline ≥20
Diastolic BP (mm Hg) (Supine or sitting – forearm at heart level)	≤50 and decrease from baseline ≥10	≥90 and increase from baseline ≥10
Pulse (bpm) (Supine or sitting)	<50 and decrease from baseline ≥15	>100 and increase from baseline ≥15
Weight (kg) (Consistent clothing and timing in relationship to meals and voiding)	(Loss) decrease ≥7%	(Gain) increase ≥7%

5.13.5. Immunogenicity

The frequency and percentage of subjects with preexisting antidrug antibody (ADA) and with treatment-emergent ADAs (TEADA) to LY3375880 will be tabulated.

For subjects who are ADA negative at baseline, TEADAs are defined as those with a titer 2-fold (1 dilution) greater than the minimum required dilution of the assay. For subjects who are ADA positive at baseline, TEADAs are defined as those with a 4-fold (2 dilution) increase in titer compared to baseline. For subjects with TEADA, the time to first TEADA result, titer evolution over time, and the distribution of maximum titers will be described.

All TEADAs will be listed.

5.13.6. Special Safety Topics, including Adverse Events of Special Interest

In addition to general safety parameters, safety information on specific topics of special interest will also be presented. Additional special safety topics may be added as warranted. The topics outlined in this section include the protocol-specified AESI.

In general, for topics regarding safety in special groups and circumstances, patient profiles and/or patient listings, where applicable, will be provided when needed to allow medical review of the time course of cases/events, related parameters, patient demographics, study drug treatment and meaningful concomitant medication use. In addition to the safety topics for which provision or review of patient data is specified, these will be provided when summary data are insufficient to permit adequate understanding of the safety topic.

5.13.6.1. Abnormal Hepatic Tests

Analyses for abnormal hepatic tests will involve 4 laboratory analytes: ALT, AST, total bilirubin, and ALP. In addition to the analyses described in Section 5.13.3, this section describes specific analyses for this topic.

The central laboratory reference ranges (CLRM reference ranges) will be used for these laboratory assessments (ALT, AST, total bilirubin, and ALP).

Analyses for change from baseline to last observation, change from the baseline value to the minimum value during the Induction/Maintenance Period, change from the baseline value to the maximum value during the Induction/Maintenance Period, and treatment-emergent high or low laboratory results at any time are described in Section 6.13.3.

The patients with the following abnormal elevations in hepatic laboratory tests at any time will be listed for both the Induction and Maintenance periods:

- ALT measurement ≥3 times, ≥5 times and ≥10 times the central laboratory ULN.
- AST measurement ≥3 times, ≥5 times and ≥10 times the central laboratory ULN.
- Total bilirubin measurement ≥2 times the central laboratory ULN.
- ALP measurement ≥1.5 times the central laboratory ULN.

5.13.6.2. Renal Function Effects

Effects on renal function will be assessed through analysis of elevated creatinine.

The CTCAE will be applied for laboratory tests related to renal effects (Table FCAB.6.4 Common Terminology Criteria for Adverse Events (CTCAE) Related to Renal Effects 4). This CTCAE grading scheme is consistent with both Version 3.0 and Version 4.03 of the CTCAE guidelines.

Shift tables will show the number and percentage of subjects from baseline to maximum during the Induction and Maintenance Periods, with baseline depicted by highest grade during the baseline period. A shift table summary displaying the number and percentage of subjects with

maximum postbaseline results will be presented by treatment group within the following categories:

- Decreased; postbaseline category < baseline category
- Increased; postbaseline category > baseline category
- Same; postbaseline category = baseline category.

Treatment-emergent lab abnormalities related to elevated creatinine occurring at any time during the Dosing Period will be tabulated using the CTCAE grades shown in Table FCAB.6.4
Common Terminology Criteria for Adverse Events (CTCAE) Related to Renal Effects

. Planned and unplanned measurements will be included.

Treatment-emergence will be characterized using 5 criteria:

- any increase in postbaseline CTCAE grade from worst baseline grade
- increase to Grade 1 or above at worst postbaseline
- increase to Grade 2 or above at worst postbaseline
- increase to Grade 3 or above at worst postbaseline
- increase to Grade 4 at worst postbaseline.

Table FCAB.6.4 Common Terminology Criteria for Adverse Events (CTCAE) Related to Renal Effects

Lab Test	CTCAE Version	Grade	Criteria in SI or CN Units
Elevated creatinine	3.0	0 (normal)	≤ ULN
		1	> ULN and ≤1.5× ULN
		2	>1.5× ULN and ≤3× ULN
		3	>3× ULN and ≤6× ULN
		4	>6× ULN

Abbreviations: CN = conventional (US); CTCAE = Common Terminology Criteria for Adverse Events; SI = Système International; ULN = upper limit of normal.

5.13.6.3. Infections

Infections will be defined using the PTs from the MedDRA Infections and Infestations SOC, with additional terms from the Investigations SOC being used in selected instances, as described below.

Treatment-emergent infections will be analyzed according to various groups of infectious events including:

- all infections
  - all PTs in the Infections and Infestations SOC,
- serious infections
  - all PTs in the Infections and Infestations SOC that are SAEs,
- infections that require therapeutic intervention (antibiotics, antivirals, antifungals, and so on)
  - all PTs in the Infections and Infestations SOC for which there is an antimicrobial concomitant medication associated with that event for that subject,

- herpes zoster
    - specific Lilly-defined PTs from the Herpes Viral Infections high-level term (HLT) in the Infections and Infestations SOC, shown in Appendix 1,
- tuberculosis
    - specific Lilly-defined PTs from the Tuberculous Infections HLT and the Investigations SOC, shown in Appendix 2
- viral hepatitis
    - all PTs from the Hepatitis Viral Infections HLT (HLT code 10057212) in the Infections and Infestations SOC.

For each infection event, the frequency for each PT will be provided, ordered by decreasing frequency in the LY3375880 highest dose group by infection group: all infections, serious infections (overall and on each approach to identifying SAEs), infections that require therapeutic intervention, herpes zoster infections, tuberculosis, and viral hepatitis.

In addition to the incidence of infectious AEs by MedDRA PT as described above, the number and percentage of subjects with treatment-emergent infectious AEs by treatment group will be summarized and listed.

Potential Opportunistic Infections:

Potential opportunistic infections (POIs) will be identified according to 2 different approaches.

POIs are identified from TEAEs based on a Lilly-defined list of MedDRA PTs, shown in the Appendix 3. These PTs are a subset of terms from the Infections and Infestations SOC.

For the POIs identified from MedDRA PTs, the number and percentage of subjects overall and for each specific PT will be summarized by treatment group, with specific event terms ordered by decreasing frequency in the LY3375880 highest dose.

Association of Infections with Lymphopenia or Neutropenia:

The relationship between the occurrence of lymphopenia and neutropenia with the occurrence of infections will be evaluated based on case reviews. Inferential analyses and/or graphical displays may be conducted if warranted. Infection outcomes considered for this analysis are any infection, any serious infection, infections that require therapeutic intervention (antibiotics, antivirals, antifungals, and so on), tuberculosis, herpes zoster, and viral hepatitis.

5.13.6.4. Allergic Reactions/Hypersensitivities

A search will be performed using the current MedDRA version 20.0 SMQs to search for relevant events, using the following queries:

- Anaphylactic reaction SMQ (20000021)
- Hypersensitivity SMQ (20000214)
- Angioedema SMQ (20000024)

Events that satisfy the queries will be listed, by temporal order within patient ID, and will include SOC, PT, SMQ event categorization including detail on the scope (narrow or broad), reported AE term, AE onset and end dates, severity, seriousness, outcome, etc.

5.13.6.5. Columbia Suicide Severity Rating Scale

Suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent, based on the C-SSRS, will be listed by patient and visit. Only patients that show suicidal ideation/behavior or self-injurious behavior without suicidal intent will be displayed (ie, if a patient's answers are all 'no' for the C-SSRS, then that patient will not be displayed). However, if a patient reported any suicidal ideation/ behavior or self-injurious behavior without suicidal intent at any time point then all their ideation and behavior will be displayed, even if not positive.

5.13.6.5.1. Self-Harm Supplement Form and Self-Harm Follow-up Form

The Self-Harm Supplement Form is a single question to enter the number of suicidal behavior events, possible suicide behaviors, or nonsuicidal self-injurious behaviors. If the number of behavioral events is greater than zero, it will lead to the completion of the Self-Harm Follow-Up Form. The Self-Harm Follow-Up Form is a series of questions that provides a more detailed description of the behavior cases. A listing of the responses give on the Self-Harm Follow-Up Form will be provided.

5.14. Subgroup Analyses

Subgroup analyses comparing each dose of LY3375880 to placebo will be performed on the ITT population at Week 16 for the following:

- Proportion of patients achieving IGA 0 or 1 with a 2-point improvement
- Proportion of patients achieving EASI75 Response Rate
- Proportion of patients achieving Itch NRS 4-point improvement

The following subgroups (but not necessarily limited to only these) will be categorized into disease-related characteristics and demographic characteristics will be evaluated:

- Patient Demographic and Characteristics Subgroups:
  - Gender (male, female)
  - Age group (<65, ≥65)
  - Diagnosis age group (<18 years, ≥18 years)
  - Baseline weight: (≤ median, > median)
  - Race: (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
  - Region: Japan vs. Rest of World
  - Immunoglobulin E (IgE): intrinsic(<200 kU/I) or extrinsic (≥200 kU/I)
  - Baseline allergen-specific IgE: Positive (≥ 0.35 kU/L) vs Negative (< 0.35 kU/L)
  - IL-19 Concentration: High (≥ 21 ng/L) vs. Low (< 21 ng/L)
  - Periostin: High (≥ 19 ng/mL) vs. Low (< 19 ng/mL)
  - TARC: High (≥ 70 pg/mL) vs. Low (< 70 pg/mL)

  - IL-13: High (≥ 0.7 pg/mL) vs. Low (< 0.7 pg/mL)
- Previous and Concomitant Therapy Subgroups:
  - Prior use of TCNI (yes, no)
  - Prior systemic therapy use (yes, no)
- Baseline Disease-Related Characteristics Subgroup
  - Duration of AD from diagnosis (0 to <2 years, 2 to < 5 years, ≥5 to <10 years, ≥10 to <20 years, ≥20 years)
  - Baseline disease severity (IGA score): 3, 4
  - Baseline EASI: (≤ median, > median)
  - Baseline BSA: (≤ median, > median)

Descriptive statistics will be provided for each treatment and stratum of a subgroup as outlined, regardless of sample size. The subgroup analyses for categorical outcomes will be performed using logistic regression. The model will include the categorical outcome as the dependent variable and treatment, subgroup, and treatment-by-subgroup interaction as explanatory variables. Missing data will be imputed using NRI (Section 5.4.1). The treatment-by-subgroup interaction comparing treatment groups will be tested at the 0.1 significance level. The p-value from the logistic regression model will be reported for the interaction test. Response counts and percentages will be summarized by treatment for each subgroup category. The difference in percentages and 100(1-alpha)% confidence interval (CI) of the difference in percentages using the Newcombe-Wilson without continuity correction will be reported. The p-value from the Fisher's exact test will also be produced.

Additional subgroup analyses on efficacy may be performed as deemed appropriate and necessary.

5.15. Protocol Deviations

Protocol deviations will be tracked by the clinical team, and their importance will be assessed by key team members during protocol deviation review meetings. Out of all important protocol deviations (IPDs) identified, a subset occurring during the Induction Period prior to the primary endpoint (Week 16) with the potential to affect efficacy analyses will result in exclusion from the PP population.

Potential examples of deviations include patients who receive excluded concomitant therapy, significant non-compliance with study medication (<80% of assigned doses taken, failure to take study medication and taking incorrect study medication), patients incorrectly enrolled in the study, and patients whose data are questionable due to significant site quality or compliance issues. Refer to a separate document for the important protocol deviations.

The number and percentage of patients having IPD(s) will be summarized within category and subcategory of deviation by treatment group for Period 2 using the ITT population. Individual patient listings of IPDs will be provided. A summary of reasons patients were excluded from the PPS will be provided by treatment group.

5.16. Interim Analyses and Data Monitoring

The assessment committee is authorized to evaluate unblinded interim efficacy and safety analyses. Study sites will receive information about interim results ONLY if they need to know for the safety of their patients.

Unblinding details are given in a separated unblinding plan.

For the Induction Period, at least one interim analyses is planned. The first interim will take place when approximately 40% to 60% patients have completed the induction phase (or discontinued induction treatment). This interim analysis will be conducted for internal decision making to trigger planning activities for future studies associated with LY3375880. The study will not be stopped early for efficacy; thus, no adjustment of type I error will be performed. The pharmacokinetic/pharmacodynamics (PK/PD) data will also be reviewed as part of the interim analysis to initiate modeling development processes. Based on emerging data, additional interim analyses may be conducted by the AC to review unblinded safety and efficacy data.

The final database will be locked when all patients complete their Week 60 (Visit 801) period or their early termination visit.

Details of the planned interim data analyses and the assessment committee data review process are included in an assessment committee charter.

5.17. Annual Report Analyses

Annual report analyses, such as the Development Update Safety Report (DSUR), will be documented in a separate document.

5.18. Clinical Trial Registry Analyses

Additional analyses will be performed for the purpose of fulfilling the Clinical Trial Registry (CTR) requirements.

Analyses provided for the CTR requirements include a summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AE are summarized: by treatment group, by MedDRA PT.

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
  - the number of participants at risk of an event
  - the number of participants who experienced each event term
  - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/patients in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).
- AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

Similar methods will be used to satisfy the European Clinical Trials Database (EudraCT) requirements.

6. References

Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology*. 2015;230(1):27-33.

Charman CR, Venn AJ, Williams, HC. The Patient-Oriented Eczema Measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. *Arch Dermatol.* 2004;140:1513-1519.

[CTCAE] Common Terminology Criteria for Adverse Events, Cancer Therapy Evaluation Program, Version 3.0, DCTD, NCI, NIH, DHHS, March 31, 2003. Available at: https://ctep.cancer.gov/protocoldevelopment/electronic_applications/docs/ctcaev3.pdf. Accessed 25 July 2017.

[CTCAE] Common Terminology Criteria for Adverse Events, Version 4.03, DHHS NIH NCI, NIH Publication No. 09-5410, June 14, 2010. Available at: https://www.eortc.be/services/doc/ctc/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf. Accessed 25 July 2017.

Gillings D, Koch G. The application of the principle of intention-to-treat to the analysis of clinical trials. *Drug Information Journal.* 1991;25(3):411-424.

Hanifin JM, Thurston M, Omoto M, Cherill R, Tofte SJ, Graeber M. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. *Exp Dermatol*. 2001;10(1):11-18.

Hongbo Y, Thomas CL, Harrison MA, Salek MS, Finlay AY. Translating the science of quality of life into practice: What do dermatology life quality index scores mean? *J Invest Dermatol*. 2005;125(4):659-664.

[ICH] International Conference on Harmonisation. Harmonised Tripartite Guideline: Clinical safety data management: definitions and standards for expedited reporting. E2A. 1994;Step 4. Available at: http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html. Accessed 21 July 2017.

Khilji FA, Gonzalez M, Finlay AY. Clinical meaning of change in Dermatology Life Quality Index scores. *Br J Dermatol*. 2002;147(suppl 62):25-54.

Kimball AB, Naegeli AN, Edson-Heredia E, Lin CY, Gaich C, Nikai E, Yosipovitch G. Psychometric properties of the Itch Numeric Rating Scale in patients with moderate-to-severe plaque psoriasis. *Br J Dermatol.* 2016;175(1):157-162.

Kunz B, Oranje A, Labrèze, Stadler JF, Ring J, Taïeb A, Clinical Validation and Guidelines for the SCORAD Index: Consensus Report of the European Task Force Atopic Dermatitis. *Dermatology*. 1997;195(1):10-19.

Naegeli AN, Flood E, Tucker J, Devlen J, Edson-Heredia E. The Worst Itch Numeric Rating Scale for patients with moderate to severe plaque psoriasis or psoriatic arthritis. *Int J Dermatol.* 2015;54(6):715-722.

[NCEP] National Cholesterol Education Program. Third Report of the NCEP Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment

Panel [ATP] III) Final Report, National Cholesterol Education Program, National Heart, Lung, and Blood Institute, National Institutes of Health, NIH Publication No. 02-5215, September 2002. *Circulation*. 2002;106(25):3143-421.

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353-365.

Schram ME, Spuls PI, Leeflang MMG, Lindeboom R, Bos JD, Schmitt J. EASI, (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. *Allergy.* 2012;67(1):99-106.

Stalder J, Taieb A, Atherton D, Bieber P, Bonifazi E, Broberg A, Calza A, Coleman R, De Prost Y, Stalder J, Gelmetti C, Cuannetti A, Harper J, Kunz B, Lachapelle J, Langeland T, Lever R, Oranje A, Oueille-Roussel C, Revuz J, Ring J, Roujeau J, Saurat J, Song M, Tennstedt D, Van Neste D, Vieluf D, Poncet M. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. *Dermatology*. 1993;186(1):23-31.

US National Archives and Records Administration. Code of Federal Regulations (CFR). Investigational New Drug Application (IND) safety reporting. 2010;Title 21:Section 312.32. Available at: https://www.ecfr.gov/cgi-bin/text-idx?SID=fd87d5f6fb00b95672dbe924a6d3b2f0&mc=true&node=se21.5.312_132&rgn=div8. Accessed 21 July 2017.

Winthrop KL, Novosad SA, Baddley JW, Calabrese L, Chiller T, Polgreen P, Bartalesi F, Lipman M, Mariette X, Lortholary O, Weinblatt ME, Saag M, Smolen J. Opportunistic infections and biologic therapies in immune-mediated inflammatory diseases: consensus recommendations for infection reporting during clinical trials and postmarketing surveillance. *Ann Rheum Dis*. 2015;74(12):2107-2116.

7. Appendices

Appendix 1. Lilly-Defined MedDRA Preferred Terms for Herpes Zoster

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Herpes zoster	10019974
Ophthalmic herpes zoster	10030865
Herpes zoster infection neurological	10061208
Herpes zoster oticus	10063491
Herpes zoster disseminated	10065038
Genital herpes zoster	10072210
Herpes zoster pharyngitis	10074245
Herpes zoster meningoencephalitis	10074248
Herpes zoster meningomyelitis	10074251
Herpes zoster meningitis	10074259
Herpes zoster cutaneous disseminated	10074297
Varicella zoster virus infection	10075611

Appendix 2. Lilly-Defined MedDRA Preferred Terms for Tuberculosis

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Adrenal gland tuberculosis	10001358
Bone tuberculosis	10056377
Bovine tuberculosis	10006049
Choroid tubercles	10008779
Congenital tuberculosis	10010657
Conjunctivitis tuberculous	10010754
Cutaneous tuberculosis	10011684
Disseminated tuberculosis	10013453
Ear tuberculosis	10014027
Epididymitis tuberculous	10015004
Erythema induratum	10015213
Extrapulmonary tuberculosis	10064445
Female genital tract tuberculosis	10061150
Immune reconstitution inflammatory syndrome associated tuberculosis	10072797
Interferon gamma release assay positive	10072866
Intestinal tuberculosis	10075268
Joint tuberculosis	10056367
Latent tuberculosis	10065048
Lupus vulgaris	10025143
Lymph node tuberculosis	10025183
Male genital tract tuberculosis	10061234
Meningitis tuberculosis	10027259
Mycobacterium tuberculosis complex test positive	10070325
Oesophogeal tuberculosis	10030200
Pericarditis tuberculous	10055069
Peritoneal tuberculosis	10053583
Prostatitis tuberculous	10064743
Pulmonary tuberculoma	10066927
Pulmonary tuberculosis	10037440
Renal tuberculosis	10038534
Salpingitis tuberculous	10039463
Silicotuberculosis	10068876
Spleen tuberculosis	10041640
Thyroid tuberculosis	10043774
Tuberculin test positive	10044728
Tuberculoma of central nervous system	10052883
Tuberculosis	10044755
Tuberculous abscess central nervous system	10052884
Tuberculosis bladder	10044758
Tuberculosis gastrointestinal	10061390

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Tuberculosis liver	10058120
Tuberculosis of central nervous system	10061391
Tuberculosis of eye	10044819
Tuberculosis of genitourinary system	10044828
Tuberculosis of intrathoracic lymph nodes	10044846
Tuberculosis of peripheral lymph nodes	10044965
Tuberculosis ureter	10045026
Tuberculous endometritis	10071559
Tuberculous laryngitis	10045072
Tuberculous pleurisy	10045104
Tuberculous tenosynovitis	10059161

Appendix 3. Lilly-Defined MedDRA Preferred Terms for Potential Opportunistic Infections

Preferred Term (MedDRA Version 18.0)	Preferred Term Code
Candida pneumonia	10053158
Respiratory moniliasis	10038705
Gastrointestinal candidiasis	10017938
Oesophageal candidiasis	10030154
Coccidioides encephalitis	10054214
Coccidioidomycosis	10009825
Cutaneous coccidioidomycosis	10068747
Meningitis coccidioides	10027207
Cryptococcal cutaneous infection	10054216
Cryptococcal fungaemia	10067112
Disseminated cryptococcosis	10013439
Gastroenteritis cryptococcal	10011485
Meningitis cryptococcal	10027209
Neurocryptococcosis	10068368
Biliary tract infection cryptosporidial	10067319
Gastroenteritis cryptosporidial	10017899
Cytomegalovirus colitis	10048983
Cytomegalovirus duodenitis	10049014
Cytomegalovirus enteritis	10049074
Cytomegalovirus enterocolitis	10049015
Cytomegalovirus gastritis	10049016
Cytomegalovirus gastroenteritis	10051349
Cytomegalovirus gastrointestinal infection	10052817
Cytomegalovirus gastrointestinal ulcer	10075619
Cytomegalovirus hepatitis	10011830
Cytomegalovirus mucocutaneous ulcer	10065036
Cytomegalovirus myelomeningoradiculitis	10065621
Cytomegalovirus myocarditis	10056261
Cytomegalovirus oesophagitis	10049018
Cytomegalovirus pancreatitis	10049566
Cytomegalovirus pericarditis	10056721
Cytomegalovirus proctocolitis	10049019
Cytomegalovirus urinary tract infection	10051350
Disseminated cytomegaloviral infection	10049075
Encephalitis cytomegalovirus	10014586