Statistical Analysis Plan for Study I1F-MC-RHBQ

A Multicenter, Randomized, Double-Blind Study Comparing the Efficacy and Safety of Ixekizumab Versus Placebo in Patients with Moderate-to-Severe Genital Psoriasis

NCT02718898

Approval Date: 08-Mar-2017

1. Statistical Analysis Plan: I1F-MC-RHBQ (b): A Multicenter, Randomized, Double-Blind Study Comparing the Efficacy and Safety of Ixekizumab Versus Placebo in Patients with Moderate-to-Severe Genital Psoriasis

ixekizumab (LY2439821) Genital Psoriasis

Study I1F-MC-RHBQ is a Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study examining the efficacy and safety of ixekizumab 80 mg Q2W (subcutaneous [SC]) as compared to placebo SC in patients with moderate-to-severe genital psoriasis, during a double-blind, 12-week Blinded Treatment Period followed by a 40-week Open-Label Treatment Period.

Eli Lilly and Company
Indianapolis, Indiana USA 46285

Protocol I1F-MC-RHBQ
Phase 3

Statistical Analysis Plan Version 1 electronically signed and approved by Lilly: 27 April 2016
Statistical Analysis Plan Version 2 electronically signed and approved by Lilly on the data provided below.

Approval Date: 08-Mar-2017 GMT

2. Table of Contents

Section **Page**

Table of Contents

Table **Page**

Table of Contents

Figure **Page**

Table of Contents

Appendix **Page**

3. Revision History

Statistical Analysis Plan (SAP) Version 1 was approved prior to unblinding. Statistical Analysis Plan Version 2 was approved prior to unblinding, a summary of changes between Version 1 and Version 2 are as follows:

Section	Summary of Changes
4.2 Secondary Objectives 6.4 Multiple Comparisons/Multiplicity 6.10 Efficacy Analyses 6.11 Analysis methodology for primary and major secondary outcomes	Added overall sPGA into major secondary objectives. Added SFQ Item 2 into major secondary objectives. Removed GPSIS Sexual Activity Avoidance Subscale from the major secondary objectives.
4.2 Secondary Objectives Table 6.5	Added change in SF-36 domain score.
6.1 General Considerations	Removed LOCF from analyses on change from baseline. Updated notations of treatment arms in Period 3.
6.1.2 Baseline Definition	Added baseline definition for GPSS, GPSIS, and SFQ, collected by e-diary. Added Multiple as a category for Race.
6.1.3.2 Secondary Analyses Method	Removed LOCF from ANCOVA analyses; removed analyses on effect size for continuous endpoint.
6.3.2 Modified Baseline Observation Carried Forward (mBOCF)	Clarified that the adverse event, mentioned in mBOCF imputation method, will include death.
6.6 Protocol Deviation	Added the category, subcategory, source and programming notes for important protocol deviations.
6.7.1 Demographics and Patient Characteristics	Removed SF-36 Total Score from baseline characteristics.
6.8.1 Previous Therapy	Added *Prior Therapy: Genital Psoriasis Entry – Topical Therapy eCRF page* into the summary of previous therapy.

Section	Summary of Changes
6.10 Efficacy Analyses	Updated derivations of GPSS item and total scores at Week 1 (Visit 3), Week 2 (Visit 4), Week 4 (Visit 5), Week 8 (Visit 6), and Week 12 (Visit 7). Updated derivations of GPSS item and total scores at each calendar week in the Blinded Treatment Period. Updated the scoring scheme for GPSIS subscale scores. Updated derivations of GPSIS subscale scores at Week 1 (Visit 3), Week 2 (Visit 4), Week 4 (Visit 5), Week 8 (Visit 6), and Week 12 (Visit 7). Updated derivations of GPSIS subscale scores at each calendar week in the Blinded Treatment Period. Updated derivations of SFQ item scores at Week 1 (Visit 3), Week 2 (Visit 4), Week 4 (Visit 5), Week 8 (Visit 6), and Week 12 (Visit 7). Updated derivations of SFQ item scores at each calendar week in the Blinded Treatment Period. Added exploratory analyses on change from baseline and/or a meaningful improvement from baseline for GPSS, GPSIS, and SFQ at each calendar week in the Blinded Treatment Period. Added exploratory analyses on GPSS Itch score for the change from baseline analyses during the first 14 days, to examine the rapid improvement from baseline in the Blinded Treatment Period. Added exploratory analyses on the association of absence/presence of genital fissure/ulcer/erosion with sPGA of genitalia at Week 12. Added exploratory analyses on the association of absence/presence of genital fissure/erosion/ulcer with measures of quality of life (SF-36 MCS, DLQI, GPSIS Subscales, and SFQ items at Week 12. Added exploratory analyses on the association of absence/presence of perianal/gluteal cleft psoriasis with measures of quality of life (SF-36 MCS, DLQI, GPSIS Subscales, and SFQ items at Week 12. Added exploratory analyses on the association of absence/presence of facial psoriasis with measures of quality of life (SF-36 MCS, DLQI, GPSIS Subscales, and SFQ items at Week 12.
6.15.1 Extent of Exposure	Added the duration of exposure to ixekizumab treatment for combined treatment period.

Section	Summary of Changes
6.15.2 Adverse Events	Added the overall summary of AEs for combined treatment period.
6.15.3.1 Special Safety Topics including Adverse Events of Special Interest	Per Ixekizumab Program Safety Analysis Plan Version 7 (IXE PSAP V7): Updated the definition on Elevated hepatic criteria by removing "with all ALP <2× ULN". Changed category of shift table for ALP from ≥1.5× ULN to >1.5× ULN to be consistent with the hepatic elevation tables. Added Medical review to identify allergic reaction/hypersensitivity, opportunistic infection, and IBD. Updated Lilly defined injection site reactions based on MedDRA V19 and removed administration site reactions since the PTs are non-specific. Added new PTs for category Squamous Cell Carcinoma based on MedDRA V19. Updated the Covance reference range to performing lab reference range. Updated Lilly-defined injection site reactions criteria. Added new PTs for category squamous cell carcinoma. Removed analyses for immunoglobulins and segmented neutrophil counts. Changed the baseline definition for neutrophil clinical recovery from Week 0 value to minimum value prior to or at Week 0. Updated the Covance reference range to performing lab reference range.
6.16.2 Immunogenicity Analyses	Updated immunogenicity definitions and terms for clarifications.

4. Study Objectives

4.1. Primary Objective

Primary Objective	Primary Endpoint
• To assess whether 80 mg ixekizumab every 2 weeks (Q2W) is superior to placebo at Week 12 in the treatment of patients with moderate-to-severe genital psoriasis as measured by static Physician's Global Assessment (sPGA) of Genitalia (0,1)	• The proportion of patients achieving sPGA of Genitalia (0,1) at Week 12

4.2. Secondary Objectives

Major Secondary Objectives	Major Secondary Endpoints
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate-to-severe genital psoriasis as measured by overall sPGA (0,1)	• The proportion of patients achieving overall sPGA (0,1) at Week 12
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate-to-severe genital psoriasis as measured by change in itch, utilizing a modified genital psoriasis itch Numeric Rating Scale (NRS) item within the Genital Psoriasis Symptoms Scale (GPSS)	• The proportion of patients with at least a 3 point improvement in genital psoriasis itch NRS within the GPSS at Week 12. This will be calculated for patients who had a baseline score of at least 3
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate-to-severe genital psoriasis as measured by limitation of frequency of sexual activity due to genital psoriasis, utilizing the SFQ item 2	• The proportion of patients whose frequency of sexual activity is never or rarely limited by genital psoriasis, an item score of 0 or 1, at Week 12. This analysis will be done among patients who had a baseline score of at least 2
Other Secondary Objectives	**Other Secondary Endpoints**
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate-to-severe genital psoriasis as measured by impact of genital psoriasis on sexual activity, utilizing the Genital Psoriasis Sexual Impact Scale (GPSIS), Sexual Activity Avoidance Subscale	• The proportion of patients whose frequency of avoiding sexual activity is either never or rarely, GPSIS Sexual Activity Avoidance Subscale score of 1 or 2, at Week 12. This will be calculated for patients who had a baseline score of at least 3
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate-to-severe genital psoriasis as measured by change in modified Genital Psoriasis Area and Severity Index (mGPASI)	• Mean change from baseline in mGPASI at Week 12

Other Secondary Objectives	Other Secondary Endpoints
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate to severe genital psoriasis as measured by the following health outcomes measures:	
○ Patient's Global Assessment of Genital Psoriasis (PatGA-Genital)	○ The proportion of patients with at least a 2-point improvement from baseline to Week 12 in PatGA-Genital. This analysis will be done among patients who had a baseline score of at least 2.
○ Dermatology Life Quality Index (DLQI) ▪ Total score ▪ Item 9 ▪ DLQI (0,1)	○ DLQI ▪ Mean change from baseline in DLQI total score at Week 12 ▪ The proportion of patients who with no sexual impairment (i.e. with a DLQI Item 9 score of 0 or 1) at Week 12 ▪ The proportions of patients achieving DLQI (0,1) at Week 12
○ Short form 36 question health survey (SF-36) ▪ physical component summary (PCS) score ▪ mental component summary (MCS) score ▪ 8 domain scores	○ SF-36 ▪ Mean change in SF-36 PCS score and the proportion of patients with at least a 2.5 point improvement at Week 12 ▪ Mean change in SF-36 MCS score and the proportion of patients with at least a 2.5 point improvement at Week 12 ▪ Mean change in SF-36 domain scores at Week 12
○ GPSS ▪ Total score ▪ Individual items	○ Mean change from baseline to Week 12 in ▪ GPSS total score ▪ and individual items
• Evaluate the incidence of anti-ixekizumab antibodies and its relationship to patient efficacy of ixekizumab at Week 12	• The proportion of patients achieving sPGA of Genitalia (0,1) at Week 12 by treatment emergent anti-drug antibody (TE-ADA) status and by neutralizing anti-drug antibody (NAb) status
• Time course of response to treatment as measured by sPGA of Genitalia (0,1)	• The proportion of patients who achieve sPGA of Genitalia (0,1) over time through Week 52
• Time course of response to treatment as measured by mGPASI	• Mean change in mGPASI over time through Week 52

4.3. Exploratory Objectives

Exploratory Objectives	Exploratory Endpoints
• To assess whether 80 mg ixekizumab every 2 weeks (Q2W) is superior to placebo at Week 12 in the treatment of genital psoriasis patients with baseline body surface area (BSA) ≥10% as measured by overall Psoriasis Area and Severity Index (PASI)	• Mean change from baseline in overall PASI at Week 12
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in genital psoriasis patients with baseline BSA ≥10% as measured by overall PASI 75/90/100	• The proportion of patients achieving overall PASI 75/90/100 at Week 12
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in patients with moderate-to-severe genital psoriasis as measured by: o static Physician Global Assessment (sPGA) of Genitalia (0) o overall sPGA (0)	• The proportion of patients achieving o sPGA of Genitalia (0) at Week 12 o overall sPGA (0) at Week 12
• To explore whether there is any impact of Fitzpatrick Skin Type on improvement measured by sPGA of Genitalia (0,1) response at Week 12	• Association between Fitzpatrick Skin Type (reported at baseline) and sPGA of Genitalia (0,1) response at Week 12
• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate to severe genital psoriasis as measured by: o GPSIS, Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale for patients who reported as being sexually active at baseline (score of "0" on GPSIS Question 1)	• To assess whether 80 mg ixekizumab Q2W is superior to placebo at Week 12 in the treatment of patients with moderate to severe genital psoriasis as measured by: o The proportion of patients whose degree of worsening is very low or none at all, or low, GPSIS Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale score of 1 or 2, at Week 12. This will be calculated for patients who reported a score of "0" on GPSIS Question 1 at baseline and had a baseline score of at least 3
o Sexual Frequency Questionnaire (SFQ) item 1	o The proportion of patients at each level of SFQ item 1
o Comprehensive Assessment of the Psoriasis Patient (CAPP) Genital sub-index	o Mean change from baseline in the total CAPP Genital sub-index score
o Touch Avoidance NRS	o Mean change from baseline in the Touch Avoidance NRS
o Short Form 36 Question Health Survey (SF-36) individual domain scores	o The proportion of patients with at least a 5 point improvement from baseline in individual domain scores

Exploratory Objectives	Exploratory Endpoints
• To explore the impact of ixekizumab versus placebo at Week 12 on the change in presence of fissure, ulcer, and/or erosion in the genital area and its association with measures of disease severity and quality of life	• The proportion of patients with presence of genital fissure, ulcer, and/or erosion and the association of presence with measures of sPGA of Genitalia (0,1) response and quality of life (SF-36 mental component summary [MCS], Dermatology Life Quality Index [DLQI], GPSIS Subscales, and SFQ items) in the subgroup of patients with fissure, ulcer, and/or erosion
• To explore the impact of ixekizumab versus placebo at Week 12 on the change in presence of psoriasis in perianal/gluteal cleft area (as indicated by the investigator in the case report form) and its association with measures of quality of life	• The proportion of patients with presence of perianal/gluteal cleft psoriasis and the association of presence with measures of quality of life (SF-36 MCS, DLQI, GPSIS Subscales, and SFQ items) in the subgroup of patients with psoriasis located in perianal area and/or gluteal cleft area
• To explore the impact of ixekizumab versus placebo at Week 12 on the change in presence of psoriasis on the face and its association with measures of quality of life	• The proportion of patients with presence of facial psoriasis and the association of presence with measures of quality of life (SF-36 MCS, DLQI, and SFQ items) in the subgroup of patients with psoriasis located on the face
• To explore long-term impact of ixekizumab on symptoms and quality of life through Week 52	• The change over time through Week 52 in DLQI, SF-36 MCS, GPSS items, GPSIS Subscales, and SFQ items
• To assess the psychometric properties (including reliability, validity, and responsiveness) of the GPSS, GPSIS, SFQ, and sPGA of Genitalia	• Test-retest reliability, construct validity, and responsiveness will be assessed by Intra-class Correlation Coefficients (ICCs), Pearson correlation/Spearman rank-based correlation coefficient, and correlations of calculated changes in scores, respectively, or as deemed appropriate
• To measure ixekizumab exposure and assess the relationship between exposure and efficacy, and exposure and immunogenicity	• Serum trough concentrations of ixekizumab • Model parameters for the exposure-response relationship between ixekizumab serum trough concentrations and efficacy endpoints • Ixekizumab serum trough concentrations associated with ADA titer subgroups
• To explore biomarkers that are predictive of response to ixekizumab treatment that may be contained in DNA, RNA, serum, or plasma samples	• Association between biomarker and ixekizumab response

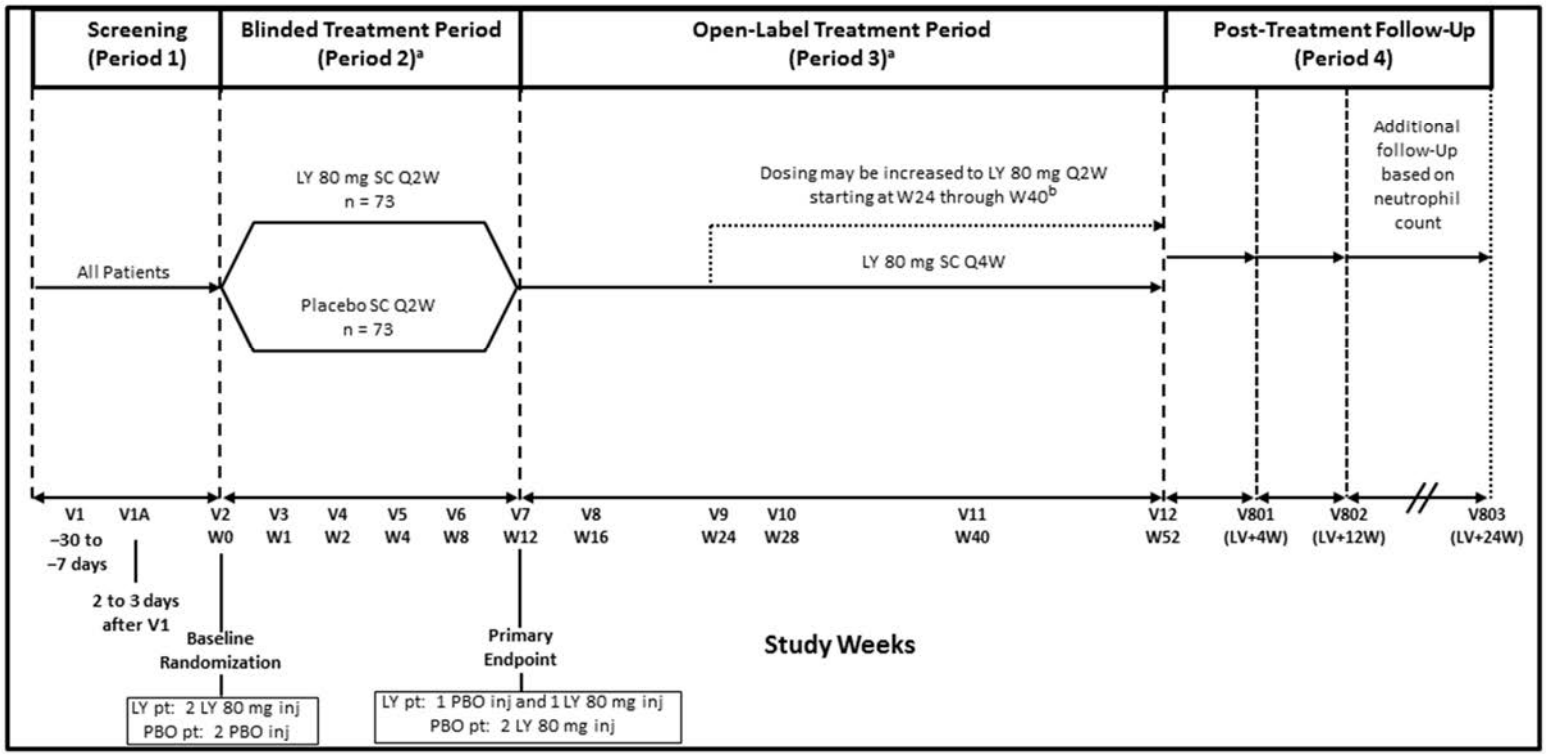
5. Study Design

5.1. Overview of Study Design

Study I1F-MC-RHBQ (RHBQ) is a Phase 3, multicenter, randomized, double-blind, placebo-controlled, parallel-group, outpatient study examining the efficacy and safety of ixekizumab as compared to placebo in patients with moderate-to-severe genital psoriasis. The study consists of 4 periods:

- **Period 1: Screening Period** (Visit 1 and Visit 1A) will assess patient eligibility and start e-diary data collection, occurring approximately 7 to 30 days prior to Period 2 (baseline; Week 0; Visit 2).
- **Period 2: Blinded Treatment Period** will occur from Week 0 (Visit 2) up to Week 12 (Visit 7). Patients will be randomized to ixekizumab or placebo in a 1:1 ratio. Two injections of ixekizumab 80 mg subcutaneous (SC) (total dose of 160 mg) or 2 injections of placebo SC, respectively, will be given at Week 0 (Visit 2). From Week 2 through Week 10 patients will receive ixekizumab 80 mg every 2 weeks (Q2W) SC or placebo Q2W SC.
- **Period 3: Open-Label Treatment Period** will occur from Week 12 (Visit 7) up to Week 52 (Visit 12). At Week 12 (Visit 7) 1 injection of ixekizumab 80 mg SC and 1 injection of placebo SC (total dose of ixekizumab 80 mg) will be given to patients who were randomized to ixekizumab in Period 2; and 2 injections of ixekizumab 80 mg SC (total dose of ixekizumab 160 mg) will be given to patients who were randomized to placebo in Period 2. During the remainder of Period 3, patients will receive ixekizumab 80 mg every 4 weeks (Q4W) dosing with an option to step-up to Q2W dosing starting at Week 24 through Week 40 (at Visit 9 [Week 24], Visit 10 [Week 28], or Visit 11 [Week 40]). This Open-label Treatment Period will allow evaluation of long-term efficacy and safety of ixekizumab through 1 year (52 weeks).
- **Period 4: Post-Treatment Follow-up** (Visit 801 through Visit 803) is for safety monitoring after treatment discontinuation for any patient receiving at least 1 dose of investigational product. Once patients complete the study treatment or discontinue study treatment early, patients will complete the Post-Treatment Follow-up (Period 4). For patients who have entered Period 4, psoriasis therapy is allowed, as determined appropriate by the investigator. This period occurs from the last treatment period visit or early termination visit (ETV) up to a minimum of 12 weeks following that visit.

Figure RHBQ.5.1 illustrates the study design.



Abbreviations: inj = injection; LV = date of last visit; LY = LY2439821 (ixekizumab); n = number of patients; pt = patient; PBO = placebo; Q2W = every 2 weeks; Q4W = every 4 weeks; SC = subcutaneous; V = visit; W = study week.

a Patients who discontinue the study for any reason and who have received at least 1 dose of investigational product will continue to Early Termination Visit before entering the Post-Treatment Follow-Up Period.

b All patients who increase dosing from 80 mg Q4W to 80 mg Q2W will remain on Q2W until completion of the study (W52 or early discontinuation) (Section 6.1).

Figure RHBQ.5.1. Illustration of study design for Clinical Protocol I1F-MC-RHBQ.

5.2. Method of Assignment to Treatment

Patients who meet all criteria for enrollment will be randomized to blinded treatment at Visit 2 (Week 0) in a 1:1 ratio to ixekizumab 80 mg Q2W or placebo. 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 double-blind investigational product to each patient. Site personnel will confirm that they have located the correct assigned investigational product package by entering a confirmation number found on the package into the IWRS.

To achieve between-group comparability for body surface area (BSA), the randomization will be stratified by BSA (1% to <10% versus ≥10%).

5.3. Determination of Sample Size

The total sample size for the study will be approximately 146 patients randomized at 1:1 ratio in the Blinded Treatment Period to ixekizumab Q2W and placebo (73 patients per treatment group).

No prior static Physician's Global Assessment (sPGA) of Genitalia (primary outcome measure) data was available at the time of planning this study to guide sample size and power calculations because no biologic drug has been evaluated in a well-controlled clinical trial to date in patients with genital psoriasis. Therefore, based on the relevance and impact of sexual impairment in genital psoriasis, data from the assessment of the Dermatology Life Quality Index (DLQI) Item 9 (How much has your skin caused any sexual difficulties?) outcomes in the ixekizumab Phase 2 and Phase 3 studies were used to calculate the sample size. Sample size was calculated assuming sexual impairment rates of 2% and 20% in the ixekizumab Q2W and placebo treatment groups, respectively. With these assumed rates, a sample size of 146 (73 per treatment group) is likely to achieve 94% power, based on a 2-sided Fisher's exact test at 0.05 level of significance.

6. A Priori Statistical Methods

6.1. General Considerations

Statistical analysis of this study will be the responsibility of Eli Lilly and Company (hereafter Lilly). The statistical analyses will be performed using SAS® Version 9.2 or higher.

Change from baseline at a particular visit will be calculated as the value at that visit minus the baseline value.

The means and medians will be reported to 1 more decimal place than the raw data recorded in the database. The standard deviation (SD) will be reported to 2 more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be 4 for any summary statistic.

Percentages will be presented to 1 decimal place. Percentages will not be presented for zero counts.

For outcome measures that are not collected at each post-baseline visit, data may exist at visits where the outcome measure was not scheduled to be collected, due to ETVs. In these situations, data from the ETVs that do not correspond to the planned collection schedule will be excluded from any mixed-effects models for repeated measures (MMRM) analysis. However, the data will still be used in other analyses, including shift analyses and other categorical analyses.

General Considerations for Analyses during Blinded Treatment Period (Period 2)

Comparisons of ixekizumab Q2W dosing versus placebo will be performed for all outcome variables in Period 2.

All confidence intervals (CIs) and statistical tests will be 2-sided unless otherwise specified. p-values which are greater than or equal to 0.001, and less than or equal to 0.999, will be presented to three decimal places. All other p-values which are less than 0.001 will be presented as <0.001, while p-values greater than 0.999 will be presented as >0.999. Confidence intervals will be presented to one more decimal place than the raw data.

General Considerations for Analyses during Open-label Treatment Period (Period 3)

Only summary statistics will be presented for the treatment groups:

- Ixekizumab 80 mg, every 2 weeks/every 4 weeks (IXE80Q2W/IXE80Q4W): Patients who received ixekizumab Q2W in the Blinded Treatment Period and ixekizumab Q4W in the Open-label Treatment Period (including those who stepped up to ixekizumab Q2W).
- Placebo/ixekizumab 80 mg, every 4 weeks (PBO/IXE80Q4W): Patients who received placebo in the Blinded Treatment Period and ixekizumab Q4W in the Open-label Treatment Period (including those who stepped up to ixekizumab Q2W).
- Overall: Unless otherwise specified, no comparisons between IXE80Q2W/IXE80Q4W and PBO/IXE80Q4W will be performed for any outcome variable in Period 3.

Data from patients who step up to ixekizumab Q2W dosing will be included in the efficacy summaries using the Open-label Treatment Population and the treatment groups defined above. The tables, figures, and listings will include footnotes indicating that patients were allowed to increase to Q2W dosing during the Open-label Treatment Period.

The number of patients who step up to Q2W dosing in Period 3 will be summarized by visit and overall using the treatment groups defined above. Patients who step up to Q2W dosing will be flagged in some efficacy and safety listings as deemed appropriate. A separate efficacy listing of sPGA of Genitalia and modified Genital Psoriasis Area and Severity Index (mGPASI) will be provided for patients who step up to Q2W dosing indicating the visits at which the patients started Q2W dosing.

6.1.1. Analysis Populations

Intent-to-treat Population: The Intent-to-treat (ITT) Population consists of all randomized patients. Even if the patients are not administered the assigned treatment, do not receive the correct treatment, or otherwise do not follow the protocol, they will be analyzed according to the treatment group to which they were assigned. Unless otherwise specified, all efficacy and health outcomes analyses during the Blinded Treatment Period will be conducted on the ITT Population.

Safety Population: The Safety Population is defined as all randomized patients who received at least 1 dose of study treatment. Patients will be analyzed according to the treatment group to which they were assigned. Unless otherwise specified, all safety analyses during the Blinded Treatment Period will be conducted on the Safety Population.

Open-label Treatment Population: The Open-label Treatment Population consists of all randomized patients who received at least 1 dose of study treatment during Period 3 and have entered the Open-label Treatment Period. Patients will be analyzed according to the treatment group to which they were assigned in Period 2, unless otherwise specified. All analyses for Period 3 (Open-label Treatment Period) will be conducted on this analysis population.

Post-Treatment Follow-Up Population: The Post-Treatment Follow-up Population consists of all randomized patients who received at least 1 dose of study treatment during Period 2 and have entered the Post-Treatment Follow-up Period. Patient data will be summarized according to the treatment group that the patient is assigned prior to entering the Post-Treatment Follow-up Period. Safety analyses for Period 4 (Post-Treatment Follow-up Period) will be conducted on this analysis population.

Table RHBQ.6.1. Treatment Group and Comparisons for Each Study Period and Analysis Population

Study Period	Analysis Population	Treatment Group	Abbreviation	Comparison
Blinded Treatment Period (Period 2)	Intent-to-treat Population; Safety Population	Ixekizumab 80 mg Q2W	IXE80Q2W	IXE80Q2W vs. PBO
		Placebo	PBO	
Open-label Treatment Period (Period 3)	Open-label Treatment Population	Ixekizumab 80 mg Q2W / Ixekizumab 80 mg Q4W	IXE80Q2W/IXE80Q4W	No comparison
		Placebo / Ixekizumab 80 mg Q4W	PBO/IXE80Q4W	
Post-Treatment Follow-up Period (Period 4)[a]	Post-Treatment Follow-up Population	Placebo	PBO	No comparison
		Ixekizumab 80 mg Q2W	IXE80Q2W	
		Ixekizumab 80 mg Q4W	IXE80Q4W	

Abbreviations: IXE = ixekizumab; PBO = placebo; Q2W = every 2 weeks; Q4W = every 4 weeks; vs. = versus.
[a] Refers to the treatment group that the patient is assigned prior to entering the Post-Treatment Follow-up Period.
Note: The treatment before the slash indicates the treatment assigned in the Blinded Treatment Period and the treatment after the slash indicates the treatment assigned in the Open-label Treatment Period.

6.1.2. Baseline Definition

Unless otherwise specified, for efficacy and health outcomes analyses in the Blinded and Open-label Treatment periods, baseline will be defined as the last available value before the first injection of the investigational product. In most cases, this will be the measurement recorded at Week 0 (Visit 2). For efficacy measures, if the patient does not take any injection, the last available value on or prior to randomization date will be used.

Unless otherwise specified for safety analyses in the Blinded Treatment Period, baseline will be defined as the last available value before the first injection of the investigational product. In most cases, this will be the measurement recorded at Week 0 (Visit 2). For safety analyses using a baseline period, the baseline period is defined as the time from Visit 1 to the date/time of the first injection.

Unless otherwise specified, the baseline for safety analyses in the Open-label Treatment Period is defined as the last non-missing assessment on or prior to Week 12 (Visit 7) and prior to the first injection of the open-label investigational product at Visit 7.

Unless otherwise specified, the baseline for safety analyses in the Post-Treatment Follow-up Period is defined as the last non-missing assessment on or prior to Week 52 (Visit 12) or ETV.

For each Genital Psoriasis Symptoms Scale (GPSS) item score and the total score, the baseline for efficacy analyses in the Blinded and Open-label Treatment Period is defined as the average of 4 or more non-missing assessments collected during the last 7 consecutive days, prior to the date of the first injection, that is, the sum of the assessments divided by the number of days on which the assessment is completed. When there are <4 non-missing assessments in the 7-day window, the window will be extended towards the date of Visit 1, until 4 daily non-missing assessments are found. However, if there are not at least 4 non-missing assessments collected prior to the date of the first injection, the baseline will be designated as missing.

For Genital Psoriasis Sexual Impact Scale (GPSIS) subscale scores and Sexual Frequency Questionnaire (SFQ) item scores, the baseline for efficacy analyses in the Blinded and Open-label Treatment Period is defined as the last non-missing assessment collected prior to the date of the first injection.

6.1.3. Analysis Methods

6.1.3.1. Primary Analysis Method

- **Categorical**
  - Treatment comparisons of all categorical efficacy and health outcome variables will be performed using a logistic regression model with treatment and BSA category (**<10%** vs **≥10% BSA at baseline**) as factors, using the NRI method. The odds ratios (OR) and the corresponding 95% CIs, will be reported.

- **Continuous**
  - The primary analyses for all continuous efficacy and health outcome variables will be performed using MMRM. The model will include treatment, baseline BSA category, baseline value, visit, treatment-by-visit interaction, and baseline-by-visit interaction as fixed factors. The covariance structure to model the within-patient errors will be unstructured. If the unstructured covariance matrix results in a lack of convergence, the heterogeneous Toeplitz covariance structure, followed by the heterogeneous autoregressive covariance structure, followed by the compound symmetry will be used. The Kenward-Roger method will be used to estimate the denominator degrees of freedom. Type III tests for the least-square (LS) means will be used for the statistical comparison; the 95% CI will also be reported. Treatment group comparisons at all visits up to Week 12 will be tested. This order is specified according to a decreasing number of covariance parameters in the structure.

6.1.3.2. Secondary Analysis Method

- Categorical
    - Secondary analysis on the categorical efficacy and health outcome variables will be conducted using a Fisher's exact test.
    - A categorical, pseudo-likelihood-based MMRM analysis for categorical repeated measures will be performed only for the primary endpoint of the sPGA of Genitalia (0,1) as analysis for estimating the percentage of patients achieving response across post-baseline visits in the Blinded Treatment Period. The model will include treatment, baseline BSA category, baseline value, visit, treatment-by-visit interaction, and baseline-by-visit interaction as fixed factors. The binomial distribution and the logit link function will be used. The restricted maximum likelihood (REML) will be used. An unstructured covariance matrix will be used to model the within-patient variance-covariance errors. The Kenward-Roger approximation will be used to estimate the denominator degrees of freedom. The Newton-Raphson with ridging optimization technique will be used to aid with convergence. The probability of response, the corresponding 2-sided 95% CI, and the p-value for the treatment group comparisons at Week 12 (Visit 7) and all other post-baseline visits will be reported.

      If the unstructured covariance matrix results in a lack of convergence, the heterogeneous Toeplitz covariance structure, followed by the heterogeneous autoregressive covariance structure, followed by the compound symmetry will be used. This order is specified according to a decreasing number of covariance parameters in the structure. The sandwich estimator (Diggle et al. 1994) for the covariance estimation will be used by specifying the EMPIRICAL option in SAS PROC MIXED. When the sandwich estimation is used, the Kenward-Roger approximation for denominator degrees of freedom cannot be used. Instead, the DDFM= BETWITHIN option will be used to estimate denominator degrees of freedom.

    - Treatment differences (Absolute Risk Reductions [ARR]), Relative Risks (RR), Numbers Needed to Treat (NNT) estimates and their corresponding 95% CIs, will also be reported.

- Continuous
    - Secondary analyses for treatment comparisons on continuous outcome variables will be conducted using an analysis of covariance (ANCOVA) model and modified baseline observation carried forward (mBOCF) imputation methods as detailed in Section 6.3.2. The ANCOVA model will include treatment, baseline BSA category, and baseline value. Type III sums of squares for the LS means will be used for the statistical comparison; the 95% CI will be reported.

6.2. Adjustments for Covariates

The randomization at the beginning of the Blinded Treatment Period (Period 2) is stratified by **baseline BSA category** (**<10%** vs **≥10% BSA at baseline**). Unless otherwise specified, all efficacy and health outcome analyses during Period 2 will include baseline BSA category in the analysis model.

In general, when an MMRM or Categorical MMRM is to be used for analyses, baseline value and baseline-by-visit interactions will be included as covariates; when an ANCOVA is to be used for analyses, baseline value will be included as a covariate.

The OBSMARGINS/OM option will be used in the LSMEANS statement to account for imbalance in the BSA categories as we expect to enroll patients in approximately 40:60 ratio from the BSA <10% and ≥10% categories.

6.3. Handling of Dropouts or Missing Data

The methods for imputation of missing data to be used in this study are in accordance with the precedent set in other Phase 3 psoriasis trials (Leonardi et al. 2008; Papp et al. 2008) and ixekizumab Phase 3 pivotal studies (I1F-MC-RHAZ [RHAZ], I1F-MC-RHBA [RHBA], and I1F-MC-RHBC [RHBC]).

The methods for imputation of missing data to be used in this study are described below.

6.3.1. Non-responder Imputation (NRI)

Analysis of categorical efficacy and health outcome variables will be based on treatment success / failure. This approach yields results numerically identical to non-responder imputation (NRI), but it is interpreted differently. Patients will be considered treatment failures if they do not meet the clinical response criteria or have missing clinical response data at a particular time point of analysis.

With NRI, as the acronym implies, there is explicit imputation of missing sPGA of Genitalia (0,1) outcomes. With treatment success / failure, discontinuation of study medication is considered a treatment failure because if patients cannot adhere to the medication they will not have sustained benefit from it. Therefore, every patient will have an observation for treatment success / failure and there will be no missing data for this estimand and hence inferences will not depend on missing data assumptions. These attributes also apply to other endpoints involving use of NRI. Randomized patients without at least 1 post-baseline observation will also be defined as non-responders for the NRI analysis.

6.3.2. Modified Baseline Observation Carried Forward (mBOCF)

An mBOCF analysis will be performed on the continuous efficacy and health outcome variables defined as "major" or "other" secondary outcome variables in Section 4.3. For patients discontinuing investigational product due to an adverse event (AE), including death, the baseline observation will be carried forward to the corresponding primary endpoint for evaluation. For patients discontinuing investigational product for any other reason, the last non-missing post-

baseline observation before discontinuation will be carried forward to the corresponding time point of evaluation. Randomized patients without at least 1 post-baseline observation will not be included for evaluation with the exception of patients discontinuing study treatment due to an AE (including death).

6.4. Multiple Comparisons/Multiplicity

A multiple testing strategy for the primary and the major secondary endpoint will be implemented to control the family-wise Type I error rate at a 2-sided α level of 0.05. The primary and major secondary endpoints will be sequentially tested in the following order to compare ixekizumab Q2W versus placebo, using the primary analysis method specified in Section 6.1.3.1.

1. Primary – Proportion of patients achieving sPGA of Genitalia (0,1) at Week 12
2. Major Secondary #1 – Proportion of patients achieving an overall sPGA (0,1) at Week 12
3. Major Secondary #2 – Proportion of patients with at least a 3 point improvement in genital psoriasis itch Numeric Rating Scale (NRS) within the GPSS at Week 12. This will be calculated for patients who had baseline score of at least 3.
4. Major Secondary #3 – Proportion of patients achieving a SFQ Item #2 score of 0 or 1, at Week 12. This will be calculated for patients who had baseline score of at least 2.

The primary endpoint will be tested at 2-sided $\alpha = 0.05$. If the test for primary endpoint is significant, then the test for the major secondary endpoint #1 will be performed. If the test for major secondary endpoint #1 is significant, then the test for major secondary endpoint #2 will be performed. If a test is not significant, all subsequent tests will be considered not significant.

There will be no adjustment for multiple comparisons for any other analyses.

6.5. Patient Disposition

A detailed description of patient disposition will be provided and the extent of their participation in the study will be reported.

Patient disposition will be listed and summarized with reasons for discontinuation from the study treatment and the study for each treatment period using the ITT, Open-label Treatment or Post-Treatment Follow-up populations.

The reasons for discontinuation from study treatment during the Blinded Treatment Period (Period 2) will be compared between treatment groups in the ITT Population using Fisher's exact test.

Time to study treatment discontinuation due to any reason (in weeks) in the Blinded Treatment Period will be summarized by treatment group and graphically presented using Kaplan-Meier technique. The log-rank test will be used to compare time to study treatment discontinuation between treatment groups. Time to study treatment discontinuation will be calculated as:

$$\frac{\text{Date of study treatment discontinuation} - \text{Date of first dose} + 1}{7}$$

If the date of first dose is missing, the date of randomization will be used. Patients completing the study treatment will be censored at the date of completion.

Patient allocation by region, country, and center/site will be summarized with number of patients who entered the study, number of ITT patients for each dosing regimen, number of patients discontinued from study treatment, and number of patients discontinued from study.

6.6. Protocol Deviations

Protocol deviations will be identified throughout the study. Important protocol deviations are defined as those deviations from the protocol likely to have a significant impact on the completeness, accuracy, and/or reliability of the study data or that may significantly affect a subject's rights, safety, or well-being.

The number and percentage of patients having important protocol deviation(s) will be summarized by treatment and category of deviations (as presented in Table RHBQ.6.2) for the Blinded Treatment Period (Period 2) using the ITT population.

A by-patient listing of important protocol deviations will be provided.

Table RHBQ.6.2. Identifications of Important Protocol Deviations

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
Failed to meet study inclusion criteria but was enrolled into the study		
[1] No confirmed diagnosis of chronic plaque Ps of at least 6 months prior to Visit 2	Monitor and Stats	Either from monitor's list, or, If date of diagnosis of psoriasis is missing or less than 6 months prior to Visit 2
[2] Not a candidate for phototherapy or systemic therapy	Monitor	From monitor's list
[3] sPGA score <3 at Visit 1 or Visit 2	Monitor and Stats	Either from monitor's list, or, If sPGA score <3 at Visit 1 or Visit 2 or any missing
[4] sPGA of Genitalia score <3 at Visit 1 or Visit 2	Monitor and Stats	Either from monitor's list, or, If sPGA of Genitalia score <3 at Visit 1 or Visit 2 or any missing
[5] Have BSA involvement of <1% at Visit1 or Visit 2	Monitor and Stats	Either from monitor's list, or, If <1% BSA involvement or any missing at Visit 1 or Visit 2
[6] Do not have confirmation of plaque psoriasis in a non-genital	Monitor	From monitor's list

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
area at Visit1 or Visit 2		
[7] Have not failed to respond to or not intolerant of, at least 1 topical therapy (corticosteroids, calcineurin inhibitors, vitamin D analogs) used for treatment of Ps affecting the genital area	Monitor and Stats	Either from monitor's list, or, If none of the medications, i.e. corticosteroids, calcineurin inhibitors, vitamin D analogs, were used and/or patient did not fail or was not intolerant of, as recorded in CRF page, Prior Therapy: Genital Psoriasis Entry – Topical Therapy.
[8] Age <18 years	Monitor	From monitor's list
[9a] Female patient with positive pregnancy test at Visit 1 and prior to randomization at Visit 2	Monitor	From monitor's list
[9b] Female patient who did not agree use a reliable method of birth control, if applicable	Monitor	From monitor's list
[9c] Male did not agree to use a reliable method of birth control	Monitor	From monitor's list
[10] Improper informed consent	Monitor and Stats	Either from monitor's list, or, If patient informed consent date is after Visit 1 date
Met study exclusion criteria but was enrolled into the study		
[11] Have predominant pattern of pustular, erythrodermic, and/or guttate forms of Ps	Monitor	From monitor's list
[12] Have pustules or vesicles in the genital area	Monitor	From monitor's list
[13] Have a history of drug-induced Ps	Monitor	From monitor's list
[14] Have received systemic non-biologic psoriasis therapy or phototherapy within 4 weeks of Visit 2 or have used topical psoriasis treatment within 2 weeks of Visit 2	Monitor	From monitor's list. Stats will program to preliminarily identify potential cases as: if have had systemic non-biologic psoriasis therapy, phototherapy or topical psoriasis therapy with psoralens≤23 days prior to Visit 2, or have had topical psoriasis treatment ≤12 days prior to Visit 2. To be specified as important protocol deviation, window of 5 days is applied for 4

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
		weeks and window of 2 days for 2 weeks. Note: The medication list will be provided by Lilly medical in a separate file.
[15] Excessive sun exposure or use of tanning booths within the 4 weeks prior to Visit 2	Monitor	From monitor's list
[16] Concurrent/recent use of biologic agent within periods prior to baseline: ETN <28 days; INF or ADA <60 days; GOL <90 days; UST <8 months; RIT <12 months; or other biologic agent <5 half-lives	Monitor and Stats	Either from monitor's list, or, If within the following washout periods prior to Visit 2: etanercept ≤23 days; infliximab or adalimumab ≤53 days; golimumab ≤80 days; ustekinumab ≤219 days; rituximab ≤330 days. To be specified as important protocol deviation, window of 5 days is applied for 28 days washout period, window of 7 days for 60 days washout period, window of 10 days for 90 days washout period, window of 21 days for 8 months washout period, window of 35 days for 12 months washout period, and window of 15 days for 5 months. Note: other biologic agent will be provided by Lilly medical in a separate file.
[17] Have ever received natalizumab or other agents that target alpha-4 integrin	Monitor	From monitor's list
[18] Have ever received treatment with IL-17 antagonists such as ixekizumab, secukinumab, or brodalumab	Monitor	From monitor's list
[19] Had a live vaccination or participated in a vaccine clinical study within 12 weeks prior to Visit 2, or intend to have a live vaccination during or within 12 weeks of completing study treatment	Monitor	From monitor's list
[20] Had a vaccination with BCG within 12 months prior to Visit 2, or intend to have this vaccination with BCG during or within 12 months of completing study treatment	Monitor	From monitor's list

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
[21] Have a known allergy or hypersensitivity to any biologic therapy that would, IOOI, pose an unacceptable risk to the patient if participating in this study	Monitor	From monitor's list
[22] Have current or a history of lymphoproliferative disease, signs or symptoms of lymphoproliferative disease, or have active or history of malignant disease within 5 years prior to Visit 2	Monitor	From monitor's list
[23] Had any major surgery within 8 weeks prior to Visit 2, or will require such during the study that, IOOI, would pose an unacceptable risk to the patient	Monitor	From monitor's list
[24] Significant uncontrolled respiratory, hepatic, renal, GI, endocrine, hematologic, neurologic, or neuropsychiatric disorder that would, IOOI, pose unacceptable risk to patient if in study	Monitor	From monitor's list
[25] Presence of significant uncontrolled cerebrocardiovascular that would, IOOI, pose an unacceptable risk to the patient if participating in the study	Monitor	From monitor's list
[26] Have ECG abnormalities that are considered clinically significant and would pose an unacceptable risk to the patient if participating in the study, IOOI	Monitor	From monitor's list
[27] Have uncontrolled arterial hypertension characterized by a systolic blood pressure (BP) >160 mm Hg or diastolic BP >100 mm Hg at Visit 1 or Visit 2 predose	Monitor and Stats	Either from monitor's list, or, If systolic BP >160 mm Hg or diastolic BP >100 mm Hg at Visit 1 or Visit 2 predose, or any missing Note: if multiple records at Visit 1 or Visit 2 predose, use the last observation
[28] Have had fluid overload, MI or new onset ischemic heart disease, uncompensated heart failure, or IOOI other serious cardiac disease within 12 weeks prior to Visit 2	Monitor	From monitor's list

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
[29] Recent (≤30 days) history of a suicide attempt, have a score of 3 on Item 12 (Thoughts of Death or Suicide) of the QIDS-SR16 at Visit 1 or Visit 2, or at risk for suicide	Monitor and Stats	Either from monitor's list, or, If have a score of 3 on Item 12 of the QIDS-SR16 at Visit 1 or Visit 2, or any missing
[30] Have evidence or suspicion of active or latent TB	Monitor	From monitor's list
[31] Are positive for human immunodeficiency virus serology (HIV)	Monitor and Stats	Either from monitor's list, or, If positive for HIV at Visit 1. Note: if multiple records at Visit 1, use the last observation
[32] Have evidence of or test positive for hepatitis B virus (HBV)	Monitor and Stats	Either from monitor's list, or, If test positive for HBV at Visit 1 by testing 1) positive for hepatitis B surface antigen (HBsAg+), OR 2) positive for anti-hepatitis B core antibody (HBcAb+) and are HBV DNA positive. Note: if multiple records at Visit 1, use the last observation. Patients who are HBcAb+ and HBV DNA negative can be enrolled. This would not be a protocol violation.
[33] Have evidence of or test positive for hepatitis C virus (HCV)	Monitor and Stats	Either from monitor's list, or, If test positive for HCV at Visit 1. A positive test for HCV is defined as: 1) positive for hepatitis C antibody (anti-HCVAb), and 2) positive via a confirmatory test for HCV (for example, HCV polymerase chain reaction). Note: if multiple records at Visit 1, use the last observation
[34] Have a body temperature ≥38°C (100.5°F) at Visit 2	Monitor and Stats	Either from monitor's list, or, If have a body temperature ≥38°C or missing at Visit 2
[35] Had a serious infection, hospitalization or IV antibiotics for an infection, a serious bone or joint infection, ever had an infection of an artificial joint, or are immunocompromised	Monitor	From monitor's list

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
[36] Have or had an infection typical of an immunocompromised host and/or that occurs with increased incidence in an immunocompromised host or have a known immunodeficiency	Monitor	From monitor's list
[37] Have or had a herpes zoster or any other clinically apparent varicella-zoster virus infection within 12 weeks of Visit 2	Monitor	From monitor's list
[38] Have any other active or recent infection within 4 weeks of Visit 2 that, IOOI, would pose an unacceptable risk to the patient	Monitor	From monitor's list
[39] Have, or are currently receiving treatment for, active candidiasis or tinea in the genital area	Monitor	From monitor's list
[40] Are currently enrolled in another CT involving IP or any other type of medical research judged not to be scientifically or medically compatible with this study	Monitor	From monitor's list
[41] Have previously completed or withdrawn from this study or participated in any other study with ixekizumab, or have participated in any study investigating other IL-17 antagonists	Monitor	From monitor's list
[42] Enrolled, participated, or discontinued from CT involving IP or non-approved use of drug/device within 30 days or 5 half-lives, whichever longer, or medical research incompatible with the study	Monitor	From monitor's list
[43] At Visit 1, have a neutrophil count <1.50 GI/L	Monitor	From monitor's list only, with statistical programming to preliminarily identify. Note: if multiple records at Visit 1, use the last observation; if V1 lab result is missing and V2 lab collection time is prior to first injection, then V2 lab result will be considered for evaluation of exclusion

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
		criterion.
[44] At Visit 1, have a lymphocyte count <0.80 GI/L	Monitor	From monitor's list only, with statistical programming to preliminarily identify. Note: if multiple records at Visit 1, use the last observation; if V1 lab result is missing and V2 lab collection time is prior to first injection, then V2 lab result will be considered for evaluation of exclusion criterion.
[45] At Visit 1, have a platelet count <100 GI/L	Monitor	From monitor's list only, with statistical programming to preliminarily identify. Note: if multiple records at Visit 1, use the last observation; if V1 lab result is missing and V2 lab collection time is prior to first injection, then V2 lab result will be considered for evaluation of exclusion criterion.
[46] At Visit 1, have aspartate aminotransferase (AST) or alanine aminotransferase (ALT) >2.5 times the upper limit of normal (ULN)	Monitor	From monitor's list only, with statistical programming to preliminarily identify. Note: if multiple records at Visit 1, use the last observation; if V1 lab result is missing and V2 lab collection time is prior to first injection, then V2 lab result will be considered for evaluation of exclusion criterion.
[47] At Visit 1, have a total white blood cell (WBC) count <3.00 GI/L	Monitor	From monitor's list only, with statistical programming to preliminarily identify. Note: if multiple records at Visit 1, use the last observation; if V1 lab result is missing and V2 lab collection time is prior to first injection, then V2 lab result will be considered for evaluation of exclusion criterion.
[48] At Visit 1, have hemoglobin <8.5 g/dL for male patients and <8.0 g/dL for female patients	Monitor	From monitor's list only, with statistical programming to preliminarily identify. Note: if multiple records at Visit 1, use the last observation; if V1 lab result is missing and V2 lab collection time is prior to first injection, then V2 lab

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
		result will be considered for evaluation of exclusion criterion.
[49] Have other clinical laboratory test results at Visit 1 that are outside the normal reference range for the population and are considered clinically significant	Monitor	From monitor's list
[50] Have donated >450 mL of blood within the last 4 weeks prior to Visit 1, or intend to donate blood during the course of the study	Monitor	From monitor's list
[51] Are women who are lactating or breastfeeding	Monitor	From monitor's list
[52] Have any other condition that precludes the patient from following and completing the protocol, IOOI	Monitor	From monitor's list
[53] Are investigator site personnel directly affiliated with this study and/or their immediate families	Monitor	From monitor's list
[54] Are Lilly employees or its designee or are employees of third-party organizations (TPOs) involved in the study	Monitor	From monitor's list
[55] Are unwilling or unable to comply with the use of a data collection device to directly record data from the subject	Monitor	From monitor's list
Met study discontinuation criteria but continued to receive study medication		
[D2a] Neutrophil counts <0.50 GI/L, or ≥0.50 GI/L and <1.00 GI/L based on 2 test results, or ≥=1.00 GI/L and <1.50 GI/L based on 3 test results and a concurrent infection	Monitor and Stats	Either from monitor's list, or, If a patient still receives study treatment after 10 days with confirmed segmented neutrophil counts <0.50 GI/L (defined as a test of <0.50 GI/L and a retest within 10 days still <0.50 GI/L; if no retest, use the test as the confirmed) Note: programming is set to identify the very first situation, that is, neutrophil counts <0.50 GI/L

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
[D2b] Total WBC count <2.00 GI/L	Monitor and Stats	Either from monitor's list, or, If a patient still receives study treatment after 10 days with confirmed total WBC count <2.00 GI/L (defined as a test of <2.00 GI/L and a retest within 10 days still <2.00 GI/L; if no retest, use the test as the confirmed)
[D2c] Lymphocyte count <0.50 GI/L	Monitor and Stats	Either from monitor's list, or, If a patient still receives study treatment after 10 days with confirmed lymphocyte count <0.50 GI/L (defined as a test of <0.50 GI/L and a retest within 10 days still <0.50 GI/L; if no retest, use the test as the confirmed)
[D2d] Platelet count <50 GI/L	Monitor and Stats	Either from monitor's list, or, If a patient still receives study treatment after 10 days with confirmed platelet count <50 GI/L (defined as a test of <50 GI/L and a retest within 10 days still <50 GI/L; if no retest, use the test as the confirmed)
[D3] Changes in BP (systolic BP at ≥160 mm Hg and ≥20 mm Hg increase from Visit 2; and/or diastolic BP at ≥100 mm Hg plus ≥10 mm Hg increase from Visit 2) that do not respond following intervention	Monitor	From monitor's list only; Statistical programming to preliminarily identify if systolic BP ≥160 mm Hg and change ≥20 mm Hg or diastolic BP ≥100 mm Hg and change ≥10 mm Hg at any post baseline visit
[D4] The patient experiences a severe AE, an SAE, or a clinically significant change in a laboratory value occurs that, IOOI, merits the discontinuation of the IP	Monitor	From monitor's list
[D5] Clinically significant systemic hypersensitivity reaction that does not respond to treatment or results in clinical sequelae	Monitor	From monitor's list
[D6] Patient became pregnant	Monitor	From monitor's list
[D7] Patient develops a malignancy other than up to 2 nonmelanoma skin cancers during the study	Monitor	From monitor's list
[D8] Change in disease phenotype	Monitor and Stats	Either from monitor's list, or, If patient still receives study treatment on the same day or after the date with an AE (PT): Guttate psoriasis, Pustular psoriasis, Erythrodermic psoriasis.

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
[D9] Required long-term treatment with a therapeutic regimen that has been demonstrated to be effective for the treatment of Ps	Monitor	From monitor's list
[D10] Enrolled in prohibited medical research	Monitor	From monitor's list
[D11] The investigator or attending physician decides that the patient should be withdrawn from study treatment	Monitor	From monitor's list
[D12] The patient requested to be withdrawn from study treatment	Monitor	From monitor's list
[D13] Investigator or Lilly stopped the patient participation	Monitor	From monitor's list
[D14] Patient became HBV DNA positive	Monitor	From monitor's list
Missing data		
Missing lab chemistry and hematology: missing baseline or not having at least 1 post-baseline	Stats	If missing lab chemistry and hematology baseline or not having at least 1 post-baseline
Missing QIDS total score: missing baseline or any scheduled visit prior to discontinuation visit	Stats	If missing QIDS total score at baseline or any post-baseline scheduled visit prior to discontinuation visit
Missing C-SSRS scale: missing baseline or any scheduled visit during period 2 or 3 prior to discontinuation visits, or any visit during period 4	Stats	If missing C-SSRS scale at Visit 2 or any other scheduled visit prior to discontinuation visit
Missing sPGA of Genitalia score: not having baseline or Week 12 measurement for patients who have completed week 12	Stats	If missing sPGA of Genitalia score at baseline or Week 12 for patients who have completed Week 12
Missing overall sPGA score: not having baseline or Week 12 measurement for patients who have completed week 12	Stats	If missing overall sPGA score at baseline or Week 12 for patients who have completed Week 12
Other		
Adverse event that meets serious criteria was not reported	Monitor	From monitor's list
Falsification (misrepresentation) of data	Monitor	From monitor's list
Site ERB approval was not obtained prior to first patient visit at site	Monitor	From monitor's list

Important Protocol Deviation Category/Subcategory	Source to Identify Protocol Deviation[a]	Statistical Programming Guidance
Unblinding is considered unjustified if the unblinding occurred for a patient where the patient's wellbeing was not dependent upon knowing their treatment assignment	Monitor	From monitor's list
Randomized but did not take any study medication		
Randomized but did not take any study medication	Stats	If a patient is randomized but does not take any study medication
Took incorrect study medication		
Took incorrect study medication	Stats	If IWRS study drug dispense data does not match the treatment label identifier on the Exposure eCRF page
Non-compliant with study medication regimen or over-dose		
Non-compliant with study medication regimen or over-dose	Stats	If non-compliant with study medication regimen or over-dose during the treatment period. Note: Non-compliance with study medication is defined to be missing more than 20% of expected doses or missing 2 or more consecutive doses; over-dose is defined as to take more injections at the same time point than specified in the protocol.
Used/took prohibited concomitant medication		
Used/took prohibited concomitant medication	Monitor	From monitor's list; stats will preliminarily program to identify potential cases as specified for prohibited concomitant medication. Note: Prohibited concomitant therapy definition will be provided by Lilly medical in a separate file.
Enrolled in a site with significant GCP non-compliance issue		
Enrolled in a site with significant GCP non-compliance issue	Monitor	From monitor's list
Had unqualified site personnel perform clinical safety and/or efficacy assessments		
Had unqualified site personnel perform clinical safety and/or efficacy assessments	Monitor	From monitor's list
e-Diary not dispensed to patient by site at Visit 1		
e-Diary not dispensed to patient by site at Visit 1	Monitor	From monitor's list

[a] The term "Monitor" indicates the protocol deviation will be identified by site monitors and entered into monitor's list (Global Protocol Deviations and Actions Items Report) using a spreadsheet. The spreadsheet will be exported as Study Data Tabulation Model (SDTM) DV domain.
The term "Stats" indicates the protocol deviation will be programmed based on data in clinical database by statistical programmers with the statistical programming guidance provided as the last column. The detailed programming specification will be documented in Analysis Data Model (ADaM) specification.
The terms "Monitor and Stats" indicates the protocol deviation will be a combination of monitor's list (SDTM.DV domain) and statistical programming from clinical database, the deviation will show if either source identifies.

Abbreviations: ADA = adalimumab; AE = adverse event; BCG = Bacille de Calmette et Guérin; BP = blood pressure; BSA = body surface area; CT = clinical trial; ECG = electrocardiogram; eCRF = electronic case report form; ETN = etanercept; GCP = good clinical practice; GI = gastrointestinal; GOL = golimumab; IL = interleukin; IOOI = in the opinion of investigator; INF = infliximab; IP = investigational product; V = intravenous; IWRS = interactive web-response system; MI = myocardial infarction; Ps = psoriasis; PT = preferred term; QIDS-SR16 = Quick Inventory of Depressive Symptomatology–Self-Report 16 items; RIT= rituximab; SAE = serious adverse event; sPGA = static Physician Global Assessment; TB = tuberculosis; UST = ustekinumab.

6.7. Patient Characteristics

6.7.1. Demographics and Baseline Characteristics

Demographics and baseline characteristics including clinical, and health outcome measurements will be summarized by treatment group and overall for the ITT and Open-label Treatment Populations. Refer to the definition of baseline in Section 6.1.2.

Categorical data for baseline variables will be summarized as frequency counts and percentages. Continuous data for baseline variables will be summarized in using descriptive statistics: number of observations, mean, standard deviation, minimum, maximum, and median. Treatment group comparisons will be conducted using Fisher's exact test for categorical data and an analysis of variance (ANOVA) model with treatment as a factor for continuous data.

By-patient listings of demographic and baseline characteristics, respectively, for the ITT population will be provided.

The following demographic and baseline characteristics will be summarized:

Demographics and Baseline Characteristics

- Sex (male, female)
- Age (in years)
- Age Categories: (<65, ≥65 to <75, ≥75 years) and (<40, ≥40 years)
- Weight (in kg)
- Weight Category: (<80, ≥80 to <100, ≥100 kg) and (<90 kg, ≥90 kg)
- Height (cm)
- Waist circumference (cm)

- Body Mass Index (BMI) (in kg/m^2)
- BMI category: (underweight [<18.5 kg/m^2]; normal [≥ 18.5 and <25 kg/m^2]; overweight [≥ 25 and <30 kg/m^2]; obese [≥ 30 and <40 kg/m^2]; obese class III [≥ 40 kg/m^2]
- Alcohol use: (never, current, former)
- Caffeine use: (never, current, former)
- Tobacco use: (never, current, former)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Multiple)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not applicable)
- Geographic region (US [including Puerto Rico], non-US)
- Country
- Previous non-biologic systemic therapy: (Never used, Used)
- Previous biologic therapy: (Never used, Used)
- Age at onset of Psoriasis (Ps)
- Age at onset of Genital Psoriasis
- Duration of disease (in years) –
    - Duration of Genital Ps since onset = [Date of informed consent – Date of onset of Genital Ps]/365.25
    - Duration of Ps since onset = [Date of informed consent – Date of onset of psoriasis]/365.25
    - Duration of Ps since diagnosed = [Date of Visit 2 – Diagnosis Date of psoriasis]/365.25
- Sociodemographics – marital and sexual partner status

Clinical and Health Outcome Measurements taken at Baseline

- sPGA for Genitalia
- sPGA for Genitalia category: (3, 4, 5)
- overall sPGA
- overall sPGA category: (3, 4, 5)
- mGPASI score
- overall Psoriasis Area and Severity Index (PASI) score
- overall PASI category: (<20, ≥ 20)

- Psoriatic arthritis: (yes, no)
- Nail psoriasis: (yes, no)
- Scalp psoriasis: (yes, no)
- Presence of Psoriasis on the face: (yes, no)
- Presence of Psoriasis on inframammary fold: (yes, no)
- Presence of Psoriasis on axilla: (yes, no)
- Presence of Psoriasis on pubis: (yes, no)
- Presence of Psoriasis on inguinal creases: (yes, no)
- Presence of Psoriasis on gluteal cleft: (yes, no)
- Presence of Psoriasis on perianal region: (yes, no)
- BSA (%)
- BSA category: ($<10\%$, $\geq 10\%$)
- Fitzpatrick skin type: (Types I, II, III, IV, V, VI)
- GPSS
    - Total score
    - Genital Ps Itch NRS score
    - Other individual item scores
- GPSIS subscales
    - Sexual Activity Avoidance Subscale
    - Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale
- Sexual Frequency Questionnaire (SFQ) item scores
- Patient's Global Assessment of Genital Psoriasis (PatGA-Genital) score
- Touch avoidance NRS score
- Short Form 36 Question Health Survey (SF-36)
    - Physical Component Summary (PCS) score
    - Mental Component Summary (MCS) score
    - Domain Scores:
        - Physical Functioning
        - Role Physical
        - Bodily Pain

  - General Health
  - Vitality
  - Social Functioning
  - Role Emotional
  - Mental Health
- DLQI
  - Total score
  - DLQI Item 9 score
  - DLQI total score of 0 or 1: (yes, no)
- Quick Inventory of Depressive Symptomatology–Self-Report 16 items (QIDS-SR16) total score
- QIDS-SR16 Item 12: 0, 1, 2, 3
- Columbia-Suicide Severity Rating Scale (C-SSRS) categories and composite scores as defined in Section 6.15.7
- Comprehensive Assessment of the Psoriasis Patient (CAPP) (genital sub-index)

6.7.2. Historical Illnesses and Pre-existing Conditions

Historical illnesses and pre-existing conditions will be classified using the latest version of the Medical Dictionary for Regulatory Activities (MedDRA).

Historical illness/condition is defined as the condition/event recorded on the Pre-existing Conditions and Medical History electronic case (clinical) report form (eCRF) page or on the Pre-specified Medical History eCRF page with an end date prior to the date of informed consent.

A pre-existing condition is defined as the condition/event recorded on the Pre-existing Conditions and Medical History eCRF page or on the Pre-specified Medical History eCRF page with a start date prior to the date of informed consent, and no end date (that is, the event is ongoing) or an end date on or after the date of informed consent. Notice if a pre-existing condition worsens in severity on or after the date of informed consent, it will be recorded as an AE on AE eCRF page from the date of worsening onwards.

The following summaries will be provided for the ITT Population:

- The number and percentage of patients with historical illnesses by treatment group and overall, by System Organ Class (SOC) and preferred term (PT).
- The number and percentage of patients with pre-existing conditions by treatment group and overall, by SOC and PT.
- The number and percentage of patients with pre-specified medical history (hypertension; diabetes mellitus, Type I; diabetes mellitus, Type II insulin dependent; diabetes mellitus

Type II non-insulin dependent; coronary artery disease; stroke; dyslipidemia; psoriatic arthritis; Crohn's disease; ulcerative colitis; Psoriasis) by treatment group and overall.

For a condition/event that is gender-specific (as defined by MedDRA), the denominator and computation of the percentage will include only patients from the given gender. The comparisons among treatment groups will be conducted using Fisher's exact test.

By-patient listings of historical illnesses and pre-existing conditions, respectively, for the ITT population will be provided.

6.8. Previous and Concomitant Therapy

A medication/therapy will be classified into anatomical therapeutic chemical (ATC) drug classes using the latest version of the World Health Organization (WHO) drug dictionary.

A by-patient listing of previous and concomitant therapy and a by-patient listing of previous psoriasis therapy and previous genital psoriasis therapy will be provided for the ITT Population.

6.8.1. Previous Therapy

Previous therapy is defined as the therapy that starts and ends prior to the date of first dose of study treatment in the Blinded Treatment Period (Period 2). If therapy start and/or end dates are missing or partial, the dates will be compared as far as possible with the date of first dose of study treatment in Period 2. If there is clear evidence to suggest that the therapy stopped prior to the first dose of study treatment in Period 2, the therapy will be assumed to be previous only.

The following summaries will be provided for the ITT population:

- Previous therapy (as captured in the Prior Therapy: Psoriasis eCRF page, the Prior Therapy: Genital Psoriasis Entry – Topical Therapy eCRF page, the Concomitant Therapy eCRF page and the Concomitant Therapy: Lipid Modifying Agents eCRF page) by WHO ATC Level 4 and WHO preferred term.
- Previous Psoriasis therapy (captured in the Prior Therapy: Psoriasis eCRF page) to be summarized according to type (non-biologic systemic agent, biologic agent, non-biologic non-systemic agent, phototherapy) and therapy. The previous biologic agent will be further classified as TNF-α inhibitor (includes infliximab, etanercept, adalimumab, golimumab, certolizumab pegol), interleukin (IL) 12/23 inhibitor (includes ustekinumab), IL-17 inhibitor (includes secukinumab), and other (includes efalizumab, alefacept, or other biological agent).
- The number and percentage of patients with each reason for discontinuation of previous Psoriasis therapy will be summarized by type and therapy.
- Previous Genital Psoriasis therapy captured in the Prior Therapy: Genital Psoriasis Entry – Topical Therapy eCRF page to be summarized according to therapy (Corticosteroids, Calcineurin Inhibitors, Vitamin D Analogs and Other Prescription Topicals).

- The number and percentage of patients with each reason for discontinuation of previous Genital Psoriasis therapy will be summarized by therapy. The number of prior genital psoriasis topical therapies taken by patients will also be summarized using the following categories: 0, 1, 2 and ≥3.

Treatment group comparisons for the Blinded Treatment Period will be performed on the ITT population using Fisher's exact test.

6.8.2. Concomitant Therapy

Concomitant therapy for Period 2 is defined as the therapy that starts before, on, or after the first day of study treatment of Period 2 and before the last visit date of Period 2, and continues into Period 2, that is, either no end date is present (the therapy is ongoing) or the end date is on or after the first day of study treatment of Period 2. Note that, a concomitant therapy will belong to Period 2 if the therapy starts and ends on the exact same day as the first day of study treatment of Period 2.

Concomitant therapy for the Open-label Treatment Period (Period 3) is defined as the therapy that starts before, on, or after the last visit date of Period 2 and continues into Period 3, that is, either no end date is present (the therapy is ongoing) or the end date is after the last visit date of Period 2. Concomitant therapy will belong to Period 3 if the therapy starts and ends on the exact same day as the first day of study treatment of Period 3.

Concomitant therapy for the Post-Treatment Follow-up Period (Period 4) is defined as the therapy that starts before, on, or after the last visit date of Period 3 and continues into Period 4, that is, either no end date is present (the therapy is ongoing) or the end date is after the last visit date of Period 3. Concomitant therapy will belong to Period 3 if the therapy starts and ends on the exact same day as the last visit date of Period 3.

The following summaries will be provided:

- Concomitant therapy (as captured in the Concomitant Therapy eCRF page and the Concomitant Therapy: Lipid Modifying Agents eCRF page) by WHO ATC Level 4 and WHO preferred term using the ITT population for the Blinded Treatment Period and the Open-label Treatment Period, using the Open-label Treatment Population for the Open-label Treatment Period and using the Post-Treatment Follow-up Population for the Post-Treatment Follow-Up Period.
- The number and percentage of patients taking concomitant therapy of topical product to be summarized for topical and topical steroid therapies, respectively, by WHO ATC Level 4 and WHO PT for the ITT population during the Blinded Treatment Period. The definition of concomitant topical therapy can be found in Appendix 5 of the Ixekizumab Program Safety Analysis Plan Version 7 (IXE PSAP V7). Please refer to I1F-MC-RHBQ (b) Clinical Protocol Section 6.8 for concomitant therapies permitted during the study.
- The number and percentage of patients who received premedication for allergic reaction/hypersensitivity captured in the Allergic / Hypersensitivity Reaction Follow-up

eCRF page during the Blinded Treatment Period (Period 2) will be summarized for the ITT Population.

Treatment group comparisons for the Blinded Treatment Period will be conducted on the ITT Population using Fisher's exact test.

6.9. Treatment Compliance

By-patient listings of randomization schedule and study drug dispensed (including the clinical trial [CT] Lot number), for the ITT Population will be provided.

Throughout the Blinded Treatment Period (Period 2), randomized patients will record information in a Study Drug Administration Log (captured in the Exposure eCRF page), including the date, time, and anatomical location of administration of investigational product, syringe number, who administered the investigational product, and the reason if the investigational product was not fully administered.

Treatment compliance for each patient will be calculated as:

$$Treatment\ Compliance\ \% = 100 \times \frac{Total\ number\ of\ injections\ administered}{Total\ number\ of\ injections\ prescribed}$$

- For patients who complete Period 2, the total number of injections prescribed during Period 2 will be equal to 7 (2 injections at Week 0 and 1 injection every 2 weeks from Week 2 to Week 10).
- For patients who discontinue during Period 2, the number of injections prescribed can be derived from the IWRS study drug dispense dataset.
- For patients who complete Period 3, the expected number of injections will be calculated as follows:
    - For patients who remain on Q4W dosing for the entire Open-label Treatment Period, the total number of injections prescribed during Period 3 will be equal to 11 (2 injections at Week 12 and 1 injection every 4 weeks from Week 16 to Week 48). The last dose will be administered at Week 48.
    - For patients who switch to Q2W dosing after Week 24, the total number of injections prescribed during Period 3 will be calculated as

      $$2 + \frac{(x - 12)}{4} + \frac{(50 - x)}{2}$$

      where *x* is the week at which the patient switched to Q2W dosing and *x* can assume values 24, 28 or 40. The last dose will be administered at Week 50.
- For patients who discontinue during Period 3, the number of injections prescribed can be derived from the IWRS study drug dispense dataset.

- The total number of injections administered will be derived using the response to the question "Was dose administered?" on the Exposure eCRF page.

A patient will be considered compliant overall for each study period if he/she misses no more than 20% of the expected doses, does not miss 2 consecutive doses, and does not overdose (that is, take more injections at the same time point than specified in the protocol).

Patient treatment compliance during the Blinded Treatment Period and the Open-label Treatment Period will be summarized for the ITT population and the Open-label Treatment Population, respectively. Proportions of patients compliant overall will be compared between treatment groups during Period 2 using Fisher's exact test.

A by-patient listing of study treatment administration and compliance for the ITT Population will be provided.

6.10. Efficacy Analyses

Table RHBQ.6.3 includes the description and derivation of the efficacy variables.

Table RHBQ.6.4 includes the description and derivation of the patient reported assessments and health outcome variables. An e-diary will be used to collect GPSS daily, GPSIS, and SFQ weekly, up to Week 12 (Visit 7); thereafter, patients will answer these questions only at scheduled study visits. At Week 12 (Visit 7), patients will complete these scales on site before completion of any other assessments.

Table RHBQ.6.5 provides the detailed analyses including analysis type, method and imputation, population, time point, and treatment group comparisons for primary, secondary and selected exploratory efficacy and health outcome analyses.

Table RHBQ.6.3. Description and Derivation of Efficacy Variables

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
sPGA of Genitalia	The Static Physician Global Assessment (sPGA) of Genitalia score is based on a combination of erythema and the secondary features (plaque elevation and/or scale). For the analysis of responses, the patient's psoriasis is assessed as follows: 0 = clear 1 = minimal 2 = mild 3 = moderate 4 = severe 5 = very severe	sPGA of Genitalia score	Ranges from 0 to 5: clear (0), minimal (1), mild (2), moderate (3), severe (4), or very severe (5).	Single item. Missing if the sPGA of Genitalia score is missing
		sPGA of Genitalia (0,1)	A sPGA of Genitalia assessed as either 0 or 1.	Missing if sPGA score of Genitalia score is missing
		sPGA of Genitalia (0)	A sPGA of Genitalia assessed as 0.	Missing if sPGA score of Genitalia score is missing
mGPASI	The Modified Genital Psoriasis Area and Severity Index (mGPASI) measure is the clinician's determination of the patient's psoriasis severity in the genital region (labia majora, labia minora, and perineum in females; penis, scrotum, and perineum in males) at a given time point yielding an overall score of 0 for no psoriasis to 72 for the most severe disease. The scoring index incorporates the degree of erythema (or redness (R)), induration (or thickness (T)), and scaling (S) of the genital plaques as well as erosion, fissure, and/or ulcer as a product of the genital area involved. Severity is rated for each	mGPASI score	The composite mGPASI score is calculated by summing the individual severity scores for erythema, induration and scaling / erosion, fissure, and/or ulcer, and then multiplying the sum by the area-of-involvement score as follows: mGPASI = (R + T + S)A Where, R = Redness score T = Thickness score S = Scaling / Erosion, fissure and/or ulcer score A = numerical value translation of % area of psoriatic involvement scores mGPASI scores are treated as a continuous score.	If any individual component score is missing, then the mGPASI score will not be calculated, and hence mGPASI score will be missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	characteristic on a 0 to 4 scale (0 for no involvement up to 4 for severe involvement including erosion, fissure, and/or ulcer). 0 = none 1 = slight 2 = moderate 3 = severe 4 = very severe The area of involvement for the entire genital area, excluding the inguinal area, is graded on a 0 to 6 scale (0 for no involvement; up to 6 for 90% to 100% involvement). 0 = 0% (clear) 1 = 1% – 9% 2 = 10% – 29% 3 = 30% – 49% 4 = 50% – 69% 5 = 70% – 89% 6 = 90% – 100%	Change from baseline in mGPASI	$Change\ from\ baseline = Observed\ mGPASI - Baseline\ mGPASI$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if baseline mGPASI or observed mGPASI value is missing
Overall sPGA	The overall sPGA is the physician's global assessment of the patient's psoriasis (Ps) lesions at a given time point. Plaques are assessed for induration, erythema, and scaling, and an overall rating of psoriasis severity is given using the anchors of 0 = clear 1 = minimal 2 = mild 3 = moderate 4 = severe	Overall sPGA score	Ranges from 0 to 5: clear (0), minimal (1), mild (2), moderate (3), severe (4), or very severe (5).	Single item. Missing if the overall sPGA score is missing
		Overall sPGA(0,1)	An overall sPGA assessed as either 0 or 1.	Missing if sPGA score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	5 = very severe	Overall sPGA(0)	An overall sPGA assessed as 0.	Missing if sPGA score is missing
Overall PASI	The overall Psoriasis Area and Severity Index (PASI): combines assessments of the extent of body surface involvement in 4 anatomical regions (head and neck, trunk, arms, and legs) and the severity of scaling (S), redness (R), and plaque induration/infiltration (thickness, T) in each region, yielding an overall score of 0 for no psoriasis to 72 for the most severe disease (Fredriksson and Pettersson 1978). Severity is rated for each index (R, S, T) on a 0-4 scale (0 for no involvement up to 4 for severe involvement): 0 = none 1 = slight 2 = moderate 3 = severe 4 = very severe The body is divided into four anatomical regions comprising the head (h), upper limb (u), trunk (t), and lower limb (l). In each of these areas, the fraction of total body surface area affected is graded on a 0-6 scale (0 for	Overall PASI total score	The composite overall PASI score is calculated by multiplying the sum of the individual severity scores for each area by the weighted area-of-involvement score for that respective area, and then summing the four resulting quantities as follows: Overall PASI $= 0.1(R_h + T_h + S_h)A_h + 0.2(R_u + T_u + S_u)A_u$ $+ 0.3(R_t + T_t + S_t)A_t + 0.4(R_l + T_l + S_l)A_l$ Where, R_h, R_u, R_t, R_l = redness score of plaques on the head, upper limb, trunk, and lower limb, scored 0-4 respectively; T_h, T_u, T_t, T_l = thickness score of plaques on the head, upper limb, trunk, and lower limb, scored 0-4 respectively; S_h, S_u, S_t, S_l = scaling score of plaques on the head, upper limb, trunk, and lower limb, scored 0-4 respectively; A_h, A_u, A_t, A_l = numerical value translation of % area of psoriatic involvement score for the head, upper limb, trunk, and lower limb, respectively. PASI scores are treated as a continuous score, with 0.1 increments within these values.	If any individual component score is missing, then the overall PASI total score will not be calculated, and hence overall PASI total score will be missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	no involvement; up to 6 for 90% - 100% involvement): 0 = 0% (clear) 1 = >0% to <10% 2 = 10% to <30% 3 = 30% to <50% 4 = 50% to <70% 5 = 70% to <90% 6 = 90% to 100% The various body regions are weighted to reflect their respective proportion of body surface area.	Overall PASI change from baseline	Calculated as: $Change\ from\ baseline = Observed\ PASI\ Total - Baseline\ PASI\ Total$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if baseline PASI Total or observed PASI Total score is missing
		Overall PASI percent improvement from baseline	Calculated as: $Percent\ improvement\ from\ baseline = -100 * \left[\frac{Observed\ PASI\ Total - Baseline\ PASI\ Total}{Baseline\ PASI}\right]$ A positive % change indicates improvement and a negative % change indicates deterioration of the condition.	Missing if baseline PASI Total or observed PASI Total score is missing
		PASI75	At least a 75% improvement in PASI score from baseline	Missing if baseline PASI Total or observed PASI Total score is missing
		PASI90	At least a 90% improvement in PASI score from baseline	Missing if baseline PASI Total or observed PASI Total score is missing
		PASI100	A 100% improvement in PASI score from baseline	Missing if baseline PASI Total or observed PASI Total score is missing

Table RHBQ.6.4. Description and Derivation of Patient Reported Assessments and Health Outcome Variables

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
GPSS	The Genital Psoriasis Symptoms Scale (GPSS) is a patient-administered assessment of 8 symptoms: itch, pain, discomfort, stinging, burning, redness, scaling, and cracking. Respondents are asked to answer the questions based on their psoriasis symptoms in the respective genital areas. The overall severity for each individual symptom from patient's genital psoriasis is indicated by selecting the number from an NRS of 0 to 10 that best describes the worst level of each symptom in the genital area in the past 24 hours, where 0 (= no severity) and 10 (worst imaginable severity). The symptom severity scores, ranging from 0 to 10, are the values of the selected numbers indicated by the patient on the instrument's horizontal scale. Each of the 8 individual items will receive a score of 0 to 10 and will be reported as item scores for itch, pain, discomfort, stinging, burning, redness, scaling, and cracking. In addition, a total score ranging from 0 (no genital psoriasis symptoms) to 80 (worst imaginable genital psoriasis symptoms) will be reported.	GPSS total score	The GPSS total score will be calculated by summing the individual item scores as follows: GPSS = itch NRS + pain NRS + discomfort NRS + stinging NRS + burning NRS + redness NRS + scaling NRS + cracking NRS The GPSS total score at Week 12 (Visit 7) is the average of at least 4 non-missing scores of the last 7 consecutive days. The 7-day window ends on the date of Visit 7 when the data is collected on site, and on or after the date of previous visit. The GPSS total score at Week 2 (Visit 4), Week 4 (Visit 5), and Week 8 (Visit 6), is the average of at least 4 non-missing scores of the last 7 consecutive days before the date of current visit, and on or after the date of previous visit. The GPSS total score at Week 1 (Visit 3), the 7-day window is before the date of Visit 3, and on or after the date of the first injection. Note 1: For derivations of post-baseline visit scores, when there are less than 4 non-missing scores in the 7-day window, the window will be extended one day at a time, up to 10 days, until 4 non-missing scores are found, but the window will not be extended beyond the date of previous visit. If at least 4 non-missing scores are still not found, the visit score will be designated as missing. Note 2: When 7-day duration is not present between visits, the visit score will be the average of at least 4	If any item score is missing, the GPSS total score will be missing.

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
			non-missing scores prior to current visit, and on or after previous visit. If at least 4 non-missing scores are not found, the visit score will be designated as missing. For the visit beyond Week 12 (Visit 7), the score is as being collected at that visit on site. As exploratory analyses, GPSS total score at each calendar week in the Blinded Treatment Period is examined, and defined as the average of 4 or more non-missing scores collected within the week. If at least 4 non-missing scores are not found during the calendar week, the score at that week is designated as missing. The calendar week is defined as follows: Calendar Week 1 is from study day 1 to 7; Calendar Week 2 is from study day 8 to 14; … Calendar Week 11 is from study day 70 to 77; Calendar Week 12 is from study day 78 to the date of Visit 7. Note that study day 1 is the date of the first injection during the Blinded Treatment Period.	
		GPSS change from baseline	$Change\ from\ baseline = Observed\ GPSS\ Total\ score - Baseline\ GPSS\ total\ score$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if either observed or baseline GPSS total score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
		GPSS item score	For each GPSS item, the score at each post-baseline visit in the Blinded Treatment Period and Open-Label Treatment Period, is derived in the same manner as GPSS total score. For each GPSS item, the score at each calendar week in the Blinded Treatment Period, is derived in the same manner as GPSS total score.	Missing if the GPSS item score is missing
		Change from baseline for each GPSS item	$Change\ from\ baseline = Observed\ NRS\ score - Baseline\ NRS\ score$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if either observed or baseline GPSS item score is missing
		3-point improvement in Genital Psoriasis Itch NRS	First calculate $Change\ from\ baseline = Observed\ Genital\ Psoriasis\ Itch\ NRS\ score - Baseline\ Genital\ Psoriasis\ Itch\ NRS\ score$ If change from baseline is less than or equal to –3, then the patient has experienced a 3-point improvement. If change from baseline is greater than –3, then the patient has not experienced a 3-point improvement.	Missing if either observed or baseline Genital Psoriasis Itch NRS score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
GPSIS	The GPSIS is a patient reported outcome measure to evaluate the impact of genital psoriasis symptoms on sexual activity. The GPSIS consists of 3 items that include 2 subscales: • **Sexual Activity Avoidance Subscale:** This subscale includes items 1 and 2. Item 1 asks whether the patient has been sexually active in the past week. Item 2 asks how often the patient avoided sexual activity in the past week due to Genital Ps. • **Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale:** This subscale includes item 1 and 3. Item 1 asks whether the patient has been	GPSIS Sexual Activity Avoidance Subscale score	For Item 1, if a patient selects "no due to reasons other than my genital psoriasis", they do not answer any additional questions on the scale and receive a score of 1 on the Sexual Activity Avoidance Subscale. If a patient responds "no due to my genital psoriasis" on GPSIS Item 1, the subscale score is set to "5" (equivalent to always avoid sexual activity due to genital psoriasis on item 2). For patients who are sexually active (i.e. with a score of 0 on Item 1), a response of never, rarely, sometimes, and often on item 2 receive a score of 1, 2, 3, and 4, respectively, on the Sexual Activity Avoidance Subscale. The Sexual Activity Avoidance Subscale ranges from 1 (never) to 5 (always) avoid sexual activity. The Avoidance Subscale Score at Week 12 (Visit 7) is the subscale score collected at Visit 7 on site. The Sexual Activity Avoidance Subscale Score at Week 2 (Visit 4), Week 4 (Visit 5), and Week 8 (Visit 6) is defined as the last non-missing subscale score collected before the date of the visit, and on or after the date of previous visit. The Avoidance Subscale score at Week 1 (Visit 3) is the last non-missing subscale score before the date of Visit 3 and on or after the date of the first injection. For the visit beyond Week 12 (Visit 7), the score is as being collected at that visit on site. For exploratory analyses, GPSS total score at each calendar week in Blinded Treatment Period is	The Sexual Activity Avoidance Subscale score is missing if: Item 1 is missing; or Item 1 is answered as "Yes" but Item 2 is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	sexually active in the past week. Item 3 asks the patient about his/her worsening of genital psoriasis symptoms following sexual activity. Total score is not calculated for the GPSIS.		examined, and defined as the last non-missing subscale score collected in the calendar week. The calendar week is defined as follows: Calendar Week 1 is from study day 1 to 7; Calendar Week 2 is from study day 8 to 14; … Calendar Week 11 is from study day 70 to 77; Calendar Week 12 is from study day 78 to the date of Visit 7. Note that study day 1 is the date of the first injection during the Blinded Treatment Period.	
		GPSIS Sexual Activity Avoidance Subscale score of 1 or 2	The subscale score is assessed as 1 or 2	Missing if the subscale score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
		GPSIS Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale	Those patients who report as being sexually active (i.e. with a score of 0 on Item 1) on the GPSIS item 1 are given the opportunity to respond to item 3. Item 3 asks the patient to select a response to reflect the level (degree) of worsening of genital psoriasis symptoms following sexual activity as described below: 1 = very low or not at all 2 = low 3 = moderate 4 = high 5 = very high The Impact Subscale Score, at each post-baseline visit in the Blinded Treatment Period and Open-Label Treatment Period, is derived in the same manner as Sexual Activity Avoidance Subscale Score. The Impact Subscale Score, at each calendar week in the Blinded Treatment Period, is derived in the same manner as Sexual Activity Avoidance Subscale Score.	The Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale score is missing if: Item 1 is missing; or Item 1 is answered as "Yes" (score of 0) but Item 3 is missing. Note: Patients, who report as not being sexually active due to genital psoriasis or other reasons (i.e. with a score of 1 or 2 on Item 1), are considered to be "Not Applicable" for this subscale, but will NOT be considered as missing.

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
		GPSIS Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale score of 1 or 2	The subscale score is assessed as 1 or 2.	Missing if the subscale score is missing
SFQ	The SFQ is a patient reported outcome measure to evaluate the impact of genital psoriasis symptoms on sexual frequency. It consists of 2 items that assess the impact of genital psoriasis symptoms on the frequency of sexual activity. Respondents are asked to answer the questions based on their psoriasis symptoms in the genital area. Item 1 asks how many times the patient engaged in sexual activity in the past week with response options of: 2 = none/zero 1 = once 0 = two or more Item 2 assesses how often genital psoriasis symptoms limited the frequency of sexual activity with the following response options: 0 = never 1 = rarely 2 = sometimes 3 = often 4 = always	SFQ Item 1	Item 1 score as reported in the weekly e-diaries. The score ranges from 0 to 2. Item 1 Score at Week 12 (Visit 7) is the score collected at Visit 7 on site. Item 1 Score at Week 2 (Visit 4), Week 4 (Visit 5), and Week 8 (Visit 6) is defined as the last non-missing score collected before the date of the visit, and on or after the date of previous visit. Item 1 Score is the lasting non-missing score before the date of Visit 3 and on or after the date of the first injection. For the visit beyond Week 12 (Visit 7), the score is as being collected at that visit on site.	Missing if the item 1 score is missing
		SFQ Item 2	Item 2 score as reported in the weekly e-diaries. The score ranges from 0 to 4. Item 2 Score, at each post-baseline visit in the Blinded Treatment Period and Open-Label Treatment Period, is derived in the same manner as Item 1 Score.	Missing if the item 2 score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	The individual item scores of the SFQ are reported separately. No total score is calculated for the SFQ.		For exploratory analyses, Item 2 Score at each calendar week is examined, and defined as the last non-missing subscale score collected in the calendar week. The calendar week is defined as follows: Calendar Week 1 is from study day 1 to 7; Calendar Week 2 is from study day 8 to 14; … Calendar Week 11 is from study day 70 to 77; Calendar Week 12 is from study day 78 to the date of Visit 7. Note that study day 1 is the date of the first injection during the Blinded Treatment Period.	
		SFQ Item 2 Score of 0 or 1	Item 2 score is assessed as 0 or 1.	Missing if the item 2 score is missing
PatGA-Genital	The Patient's Global Assessment of Genital Psoriasis (PatGA-Genital) is a patient-administered, single-item scale on which patients are asked to rank the severity	PatGA-Genital score	The PatGA-Genital NRS score as reported by the patient. The score ranges from 0 to 5.	Missing if the PatGA-Genital score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	of their genital psoriasis "today" by circling a number on a 0 to 5 NRS, as follows: from 0 (clear), no genital psoriasis; to 5 (severe).	2 point improvement in PatGA-Genital score	First calculate $Change\ from\ baseline = Observed\ PatGA\ Genital\ score - Baseline\ PatGA\ Genital\ score$ If change from baseline is less than or equal to –2, then the patient has experienced a 2-point improvement. If change from baseline is greater than –2, then the patient has not experienced a 2-point improvement.	Missing if the observed or baseline PatGA Genital score is missing
DLQI	The DLQI is a simple, patient-administered, 10 question, validated, quality-of-life questionnaire that covers 6 domains: 1. Symptoms and feelings 2. Daily activities 3. Leisure 4. Work and school 5. Personal relationships 6. Treatment Response categories include 0 = not at all 1 = a little 2 = a lot and 3 = very much "not relevant" responses scored as "0".	DLQI total score	A DLQI total score is calculated by summing all 10 question responses, and has a range of 0 to 30 (less to more impairment). (Finlay and Khan 1994; Basra et al. 2008).	If two or more questions are missing, the total score is missing.
		Change from baseline in DLQI total score	$Change\ from\ baseline = Observed\ DLQI\ Total\ score - Baseline\ DLQI\ Total\ score$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if the observed or baseline DLQI Total score is missing
	DLQI Item 9 asks the following: How much has your skin caused any sexual difficulties?	DLQI Item 9 score	DLQI item 9 score as reported by the patient.	Missing if the DLQI item 9 score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
		No Sexual Impairment	A DLQI Item 9 score of 0 or 1	Missing if the DLQI item 9 score is missing
	A DLQI total score of 0 to 1 is considered as having no effect on a patient's Health Related Quality of Life (HRQoL) (Khilji et al. 2002; Hongbo et al. 2005).	DLQI (0,1)	A DLQI (0,1) response is defined as a post-baseline DLQI total score of 0 or 1.	Missing if the DLQI total score is missing
	DLQI domains • **Symptoms and feelings domain** #1. How itchy, sore, painful or stinging has your skin been? #2. How embarrassed or self-conscious have you been because of your skin? • **Daily activities** #3. How much has your skin interfered with you going shopping or looking after your home or garden? #4. How much has your skin influenced the clothes you wear? • **Leisure** #5. How much has your skin affected any social or leisure activities? #6. How much has your skin make it difficult for you to do any sport? • **Work and school** #7A. Has your skin prevented you from working or studying? #7B. If No: how much has your skin been a problem at work or studying? • **Personal relationships** #8. How much has your skin created problems with your partner or any of	DLQI Domain scores	The domain scores are calculated as follows: • Symptoms and feelings (Sum of Q1 and Q2 scores) • Daily activities (Sum of Q3 and Q4 scores) • Leisure (Sum of Q5 and Q6 scores) • Work and school (Sum of Q7A and Q7B scores) • Personal relationships (Sum of Q8 and Q9 scores) • Treatment (Q10 score) A lower value indicates less impairment and a higher value indicates more impairment.	If one question in a domain is missing, that domain is missing. Note: #7B could be a valid missing while #7A is not "No." That is, #7 should be considered as one question.
		Change from baseline in a DLQI domain score	$Change\ from\ baseline = Observed\ DLQI\ domain\ score - Baseline\ DLQI\ domain\ score$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if the observed or baseline DLQI domain score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	your close friends or relatives? #9. How much has your skin caused any sexual difficulties? • **Treatment** #10. How much of a problem has the treatment for your skin been, for example by making your home messy, or by taking up time?			
SF-36	The Short-Form Health Survey (SF-36) is a 36-item, patient-completed measure designed to be a short, multipurpose assessment of health in the areas of • physical functioning • role-physical • role-emotional • bodily pain • vitality • social functioning • mental health • general health The 2 overarching composite domains are 1. Mental Component Summary (MCS) scores and 2. Physical Component Summary (PCS) The summary scores are normalized and transformed to calculate the Physical (PCS) and Mental Component (MCS) summary scores with a normative value of 50 and standard deviation of 10.	SF-36 MCS score	The details of the derivation of SF-36 MCS score are presented in Appendix 1	See details in Appendix 1
		Change from baseline in SF-36 MCS score	$Change\ from\ baseline = Observed\ SF36\ MCS\ score - Baseline\ SF36\ MCS\ score$ A positive change indicates improvement and a negative change indicates deterioration of the condition.	Missing if the observed or baseline SF-36 MCS score is missing
		SF-36 PCS score	The details of the derivation of SF-36 PCS score are presented in Appendix 1	See details in Appendix 1

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	Minimal clinically important differences (MCID) are defined as ≥2.5-point increases from baseline for SF-36 PCS and MCS. Items are answered on Likert scales of varying lengths. The SF-36 acute version will be used, which has a 1-week recall period	Change from baseline in SF-36 PCS score	$Change\ from\ baseline = Observed\ SF36\ PCS\ score - Baseline\ SF36\ PCS\ score$ A positive change indicates improvement and a negative change indicates deterioration of the condition.	Missing if the observed or baseline SF-36 PCS score is missing
		2.5 point improvement in SF-36 PCS	First calculate $Change\ from\ baseline = Observed\ SF36\ PCS\ score - Baseline\ SF36\ PCS\ score$ If change from baseline is greater than or equal to 2.5, then the patient has experienced a 2.5-point improvement. If change from baseline is less than 2.5, then the patient has not experienced a 2.5-point improvement.	Missing if the observed or baseline SF-36 PCS score is missing
		2.5 point improvement in SF-36 MCS	First calculate $Change\ from\ baseline = Observed\ SF36\ MCS\ score - Baseline\ SF36\ MCS\ score$ If change from baseline is greater than or equal to 2.5, then the patient has experienced a 2.5-point improvement. If change from baseline is less than 2.5, then the patient has not experienced a 2.5-point improvement.	Missing if the observed or baseline SF-36 MCS score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	The 8 domain scores of SF-36 are: • Physical Functioning • Role-Physical • Role-Emotional • Bodily Pain • Vitality • Social Functioning • Mental Health • General Health The domain scores range from 0-100 with higher scores indicating better HRQoL. Domain score MCID is a ≥5.0-point increase from baseline.	SF-36 Domain scores	The details of the derivation of SF-36 domain scores are presented in Appendix 1 A higher domain score indicates lesser impairment.	Missing if the SF-36 domain score is missing
		Change from baseline in SF-36 domain score	$Change\ from\ baseline = Observed\ SF36\ domain\ score - Baseline\ SF36\ domain\ score$ A positive change indicates improvement and a negative change indicates deterioration of the condition.	Missing if the observed or baseline SF-36 doscore is missing
		5 point improvement in SF-36 domain score	First calculate $Change\ from\ baseline = Observed\ domain\ score - Baseline\ domain\ score$ If change from baseline is greater than or equal to 5, then the patient has experienced a 5-point improvement. If change from baseline is less than 5, then the patient has not experienced a 5-point improvement.	Missing if the observed or baseline SF-36 domain score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
CAPP Genital Severity Index	An assessment of psoriasis severity, combining both clinician and patient evaluation to determine severity of overall plaque psoriasis as well as several sub-indices (scalp, nail, palmar-plantar, inverse, and genital). **For this study, we will assess only the genital severity sub-index.** CCI - This section contained a clinical outcome assessment questionnaire that is protected by 3rd party copyright law.	CAPP Genital Severity Index score	The genital severity score measured by the clinician ranges from 0 to 10. Each of the two patient reported outcomes is based on a visual analog scale (VAS) of 0 to 10. $CAPP\ Genital\ Severity\ Index\ score = Clinician\ reported\ genital\ severity\ score + Max[Intimacy\ VAS, Pain\ VAS]$ The CAPP Genital Severity Index score will range from 0 (clear/no impact) to 20 (severe genital psoriasis/worst imaginable pain and/or unable to be intimate at all).	The CAPP Genital severity index score will be set to missing if the clinician reported genital severity score is missing OR both of the patient reported VAS scores are missing.
		Change from baseline in CAPP Genital Severity Index score	$Change\ from\ baseline = Observed\ CAPP\ genital\ severity\ index\ score - Baseline\ CAPP\ genital\ severity\ index\ score$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if the observed or baseline CAPP Genital severity index score is missing
Touch Avoidance NRS	The Touch Avoidance (TA) Numeric Rating Scale is a self-administered, single item scale that assesses touch avoidance over the past 2 weeks due to the look or feel of the patient's	Touch Avoidance NRS score	The item is rated on scale of 0 (not at all) to 10 (very much).	Missing if the NRS score is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
	skin.	Change from baseline in Touch Avoidance NRS score	$Change\ from\ baseline = Observed\ TA\ NRS\ score - Baseline\ TA\ NRS\ score$ A negative change indicates improvement and a positive change indicates deterioration of the condition.	Missing if the observed or baseline TA NRS score is missing
Visible Ps in specific Locations	The presence of visible Psoriasis on the following locations will be noted: 1. Face 2. Inframammary fold 3. Axilla 4. Scalp 5. Nail 6. Pubis 7. Perianal region 8. Gluteal cleft and 9. Inguinal creases	Visible Ps Location - Overall	Binary responses (yes/no) will be noted to determine the presence of visible psoriasis on each of the 9 locations.	Missing if the response is missing
	The presence of visible psoriasis in the following genital areas: For women 1. Labia minora 2. Labia majora 3. Perineum For men 1. Penis (glans and/or shaft) 2. Scrotum 3. Perineum	Visible Ps Location - Genital	Binary responses (yes/no) will be noted to determine the presence of visible psoriasis in the genital areas	Missing if the response is missing

Measure	Description	Variable	Derivation	Imputation Approach if Components are Missing
Genital Ps characteristics	The presence of fissure, erosion, and ulcer in the genital area in the following genital areas: For women 1. Labia minora 2. Labia majora 3. Perineum For men 1. Penis 2. Scrotum 3. Perineum	Genital Ps characteristics	Binary responses (yes/no) will be noted to determine the presence of fissure, erosion and ulcer in the genital areas	Missing if the response is missing
Perianal / Gluteal characteristics	The presence of fissure, abscess, ulcer and skin tags in the following areas: In the perianal area • Fissure • Abscess • Ulcer • Skin tags In the gluteal cleft • Fissure	Perianal area	Binary responses (yes/no) will be noted to determine the presence of fissure, abscess, ulcer and skin tags in the perianal area	Missing if the response is missing
		Gluteal cleft	Binary responses (yes/no) will be noted to determine the presence of fissure in the gluteal cleft	Missing if the response is missing

Note: Unless otherwise specified, genital area means labia minora, labia majora and perineum for women and penis, scrotum and perineum for men.

Table RHBQ.6.5. Analysis Methods for Efficacy and Health Outcome Variables

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
sPGA of Genitalia	Proportions of patients with sPGA of Genitalia(0,1)	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	**Primary (at Week 12 only)**	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
			Week 12 and other post-baseline visits	Categorical MMRM using observed data		Secondary
	Proportions of patients with sPGA of Genitalia(0)	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Exploratory	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
Overall sPGA	Proportion of patients with Overall sPGA(0,1)	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	**Major Secondary #1 (at Week 12 only)**	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
	Proportion of patients with Overall sPGA(0)	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Exploratory	Primary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	Proportion of patients with Overall sPGA(0)	ITT Population	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Secondary
GPSS[a]	Proportion of patients with at least a 3-point improvement in Genital Ps Itch NRS score	ITT Population: Patients with a baseline score of at least 3	Week 12 and other post-baseline visits[b]	Logistic regression analysis using NRI	**Major Secondary #2 (at Week 12 only)**	Primary
			Week 12 and other post-baseline visits[b]	Fisher's exact test using NRI		Secondary
			At each calendar week[c]	Logistic regression analysis using NRI	Exploratory	Primary
			At each calendar week[c]	Fisher's exact test using NRI		Secondary
	Change from baseline in GPSS Total score and Item scores	ITT Population	Week 12 and other post-baseline visits[b]	MMRM	Other secondary	Primary
			Week 12 and other post-baseline visits[b]	ANCOVA using mBOCF		Secondary
			At each calendar week[c]	MMRM	Exploratory	Primary
	Change from baseline in Genital Ps Itch NRS daily score	ITT Population	During the first 14 days	MMRM	Exploratory	Primary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	Time to at least a 3-point improvement in Itch NRS score	ITT Population: Patients with a baseline score of at least 3	During the Blinded Treatment Period	Kaplan Meier Analyses	Exploratory	Primary
SFQ[a]	Proportion of patients achieving SFQ item 2 score of 0 or 1	ITT Population: Patients with a baseline score of at least 2	Week 12 and other post-baseline visits[b]	Logistic regression analysis using NRI	**Major Secondary #3 (at Week 12 only)**	Primary
			Week 12 and other post-baseline visits[b]	Fisher's exact test using NRI		Secondary
			At each calendar week[c]	Logistic regression analysis using NRI	Exploratory	Primary
			At each calendar week[c]	Fisher's exact test using NRI		Secondary
	Proportion of patients at each level of SFQ item 1	ITT Population	Week 12 and other post-baseline visits[b]	Fisher's exact test based on observed data	Exploratory	Primary
GPSIS[a]	Proportion of patients achieving Sexual Activity Avoidance	ITT Population: Patients with a baseline score of at least 3	Week 12 and other post-baseline visits[b]	Logistic regression analysis using NRI	**Other secondary**	Primary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	Subscale score of 1 or 2		Week 12 and other post-baseline visits[b]	Fisher's exact test using NRI		Secondary
			At each calendar week[c]	Logistic regression analysis using NRI	Exploratory	Primary
			At each calendar week[c]	Fisher's exact test using NRI		Secondary
	Proportion of patients achieving Impact of Sexual Activity on Genital Psoriasis Symptoms Subscale score of 1 or 2	ITT Population: Patients with an Item 1 score of "0" and an Item 3 score of at least 3 at baseline	Week 12 and other post-baseline visits[b]	Logistic regression analysis using NRI	Exploratory	Primary
			Week 12 and other post-baseline visits[b]	Fisher's exact test using NRI		Secondary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
			At each calendar week[c]	Logistic regression analysis using NRI	Exploratory	Primary
			At each calendar week[c]	Fisher's exact test using NRI		Secondary
mGPASI	Change from baseline in mGPASI	ITT Population	Week 12 and other post-baseline visits	MMRM	Other secondary	Primary
			Week 12 and other post-baseline visits	ANCOVA using mBOCF		Secondary
PatGA-Genital	Proportion of patients with at least a 2-point improvement in PatGA-Genital	ITT Population: Patients with a baseline score of at least 2	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Other secondary	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
DLQI	Change from baseline in DLQI total score	ITT Population	Week 12 and other post-baseline visits	MMRM	Other secondary	Primary
			Week 12 and other post-baseline visits	ANCOVA using mBOCF		Secondary
	Proportion of patients with no sexual impairment	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Other secondary	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
	Proportion of patients with DLQI(0,1)	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Other secondary	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
SF-36	Change from baseline in SF-36 PCS score	ITT Population	Week 12 and other post-baseline visits	ANCOVA using mBOCF	Other secondary	Secondary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	Proportion of patients with at least a 2.5 point improvement in SF-36 PCS	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Other secondary	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
	Change from baseline in SF-36 MCS score	ITT Population	Week 12 and other post-baseline visits	ANCOVA using mBOCF	Other secondary	Secondary
	Proportion of patients with at least a 2.5 point improvement in SF-36 MCS	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Other secondary	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
	Proportion of patients with at least a 5 point improvement in SF-36 individual domain scores	ITT Population	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Other secondary	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
Overall PASI	Change from baseline in Overall PASI score	ITT Population: Patients with BSA ≥ 10%	Week 12 and other post-baseline visits	MMRM	Exploratory	Primary
	Percentage Improvement from baseline in Overall PASI score	ITT Population: Patients with BSA ≥ 10%	Week 12 and other post-baseline visits	MMRM	Exploratory	Primary
	PASI 75	ITT Population: Patients with BSA ≥ 10%	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Exploratory	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
	PASI 90	ITT Population: Patients with BSA ≥ 10%	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Exploratory	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary
	PASI 100	ITT Population: Patients with BSA ≥ 10%	Week 12 and other post-baseline visits	Logistic regression analysis using NRI	Exploratory	Primary
			Week 12 and other post-baseline visits	Fisher's exact test using NRI		Secondary

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
CAPP	Change from baseline in CAPP genital sub-index total score	ITT Population	Week 12 and other post-baseline visits	MMRM	Exploratory	Primary
Touch Avoidance NRS	Change from baseline in Touch Avoidance NRS score	ITT Population	Week 12 and other post-baseline visits	MMRM	Exploratory	Primary
Fissure, ulcer, and/or erosion in genital area	Proportion of patients with absence of fissure	ITT Population – Patients with presence of fissure	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with absence of fissure/erosion/ulcer	ITT Population – Patients with presence of fissure, erosion, and/or ulcer at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with absence of erosion	ITT Population – Patients with presence of erosion at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
Psoriasis in perianal or gluteal cleft area	Proportion of patients with absence of psoriasis on perianal area	ITT Population – Patients with presence of psoriasis on perianal area at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with absence of psoriasis on gluteal cleft	ITT Population – Patients with presence of psoriasis on gluteal cleft at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
Psoriasis in peri-genital area	Proportion of patients with absence of psoriasis on pubis	ITT Population – Patients with presence of psoriasis on pubis at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with absence of psoriasis on inguinal creases	ITT Population – Patients with presence of psoriasis on inguinal creases at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
Psoriasis on face, nail, scalp or axilla	Proportion of patients with absence of psoriasis on face	ITT Population – Patients with presence of psoriasis on face at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with absence of psoriasis on nail	ITT Population – Patients with presence of psoriasis on nail at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with absence of psoriasis on scalp	ITT Population – Patients with presence of psoriasis on scalp at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with absence of psoriasis on axilla	ITT Population – Patients with presence of psoriasis on axilla at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
Psoriasis on inframammary fold	Proportion of patients with absence of psoriasis on inframammary fold	ITT Population – Patients with presence of psoriasis on inframammary fold at baseline	Week 12 and other post-baseline visits	Fisher's exact test using NRI	Exploratory	Primary[e]
sPGA of Genitalia and genital fissure/erosion/ulcer	Proportions of patients with absence of genital fissure, compared between two groups of patients achieving and not achieving sPGA of Genitalia (0,1) [d]	ITT Population – Patients with genital fissure at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportions of patients with absence of genital erosion, compared between two groups of patients achieving and not achieving sPGA of Genitalia (0,1) [d]	ITT Population – Patients with genital erosion at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportions of patients with absence of genital fissure/erosion/ ulcer, compared between two groups of patients achieving and not achieving sPGA of Genitalia (0,1) [d]	ITT Population – Patients with genital fissure, erosion, and/or ulcer at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
SFQ Item 1 and genital fissure/ erosion/ ulcer	Proportion of patients with SFQ Item 1 score response level, compared between two groups of patients with presence/absence of genital fissure[d]	Intent to Treat Population – Patients with genital fissure at baseline	Week 12	Fisher's exact test	Exploratory	Primary[e]
	Proportion of patients with SFQ Item 1 score response level, compared between two groups of patients with presence/absence of genital erosion[d]	Intent to Treat Population – Patients with genital erosion at baseline	Week 12	Fisher's exact test	Exploratory	Primary[e]
	Proportion of patients with SFQ Item 1 score response level, compared between two groups of patients with presence/absence of genital fissure, erosion, and /or ulcer[d]	Intent to Treat Population – Patients with genital fissure, erosion, and/or ulcer at baseline	Week 12	Fisher's exact test	Exploratory	Primary[e]
SFQ Item 1 and psoriasis in perianal/ gluteal cleft area	Proportion of patients with SFQ Item 1 score response level, compared between two groups of patients with	Intent to Treat Population – Patients with psoriasis in perianal and/or gluteal cleft area at baseline	Week 12	Fisher's exact test	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	presence/absence of psoriasis in perianal /gluteal cleft area[d]					
SFQ Item 1 and facial psoriasis	Proportion of patients with SFQ Item 1 score response level, compared between two groups of patients with presence/absence of facial psoriasis[d]	Intent to Treat Population – Patients with facial psoriasis at baseline	Week 12	Fisher's exact test	Exploratory	Primary[e]
SFQ Item 2 and genital fissure/ erosion/ ulcer	Proportion of patients with SFQ Item 2 score of 0 or 1, compared between two groups of patients with presence/ absence of genital fissure[d]	Intent to Treat Population – Patients with SFQ Item 2 score ≥ 2 and with genital fissure at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with SFQ Item 2 score of 0 or 1, compared between two groups of patients with presence/absence of genital erosion[d]	Intent to Treat Population – Patients with SFQ Item 2 score ≥ 2 and with genital erosion at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with SFQ Item 2 score of 0 or 1, compared between two	Intent to Treat Population – Patients with SFQ Item 2 score ≥ 2 and	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	groups of patients with presence/absence of genital fissure, erosion, and /or ulcer[d]	with genital fissure, erosion, and/or ulcer at baseline				
SFQ Item 2 and psoriasis in perianal/ gluteal cleft area	Proportion of patients with SFQ Item 2 score of 0 or 1, compared between two groups of patients with presence/absence of psoriasis in perianal/gluteal cleft area[d]	Intent to Treat Population – Patients with SFQ Item 2 score ≥ 2 and psoriasis in perianal and/or gluteal cleft area at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
SFQ Item 2 and facial psoriasis	Proportion of patients with SFQ Item 2 score of 0 or 1, compared between two groups of patients with presence/absence of facial psoriasis[d]	Intent to Treat Population – Patients with SFQ Item 2 score ≥ 2 and facial psoriasis at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
GPSIS Avoidance Subscale and facial psoriasis	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/ absence of facial psoriasis[d]	Intent to Treat Population – Patients with GPSIS Avoidance Subscale Score ≥3 and facial psoriasis at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
GPSIS Avoidance Subscale and genital fissure/ erosion/ ulcer	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/ absence of genital fissure[d]	ITT population – Patients with GPSIS Avoidance Subscale Score ≥3 and genital fissure at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/ absence of genital erosion[d]	ITT population – Patients with GPSIS Avoidance Subscale Score ≥3 and genital erosion at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/ absence of genital fissure/erosion/ ulcer[d]	ITT population – Patients with GPSIS Avoidance Subscale Score ≥3 and genital fissure, erosion, and/ or ulcer at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
GPSIS Avoidance Subscale and psoriasis in perianal/ gluteal cleft area	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/ absence of	ITT population – Patients with GPSIS Avoidance Subscale Score ≥3 and psoriasis in perianal and/or gluteal	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	psoriasis in perianal/gluteal cleft area[d]	cleft area at baseline				
GPSIS Impact Subscale and genital fissure/ erosion/ ulcer	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/absence of genital fissure[d]	ITT population – Patients with GPSIS Impact Subscale Score ≥3 and genital fissure at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/absence of genital erosion[d]	ITT population – Patients with GPSIS Impact Subscale Score ≥3 and genital erosion at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/absence of genital fissure/erosion/ ulcer[d]	ITT population – Patients with GPSIS Impact Subscale Score ≥3 and genital fissure, erosion, and/or ulcer at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
GPSIS Impact Subscale and psoriasis in perianal/ gluteal cleft area	Proportion of patients with a subscale score of 1 or 2, compared between two groups of patients with presence/absence of psoriasis in perianal/gluteal cleft area[d]	ITT population – Patients with GPSIS Impact Subscale Score ≥3 and psoriasis in perianal and/or gluteal cleft area at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
DLQI and genital fissure/ erosion/ ulcer	Proportion of patients with DLQI (0,1), compared between two groups of patients with presence/absence of genital fissure[d]	Intent to Treat Population – Patients with genital fissure at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with SFQ Item 2 score of 0, compared between two groups of patients with presence/absence of genital erosion[d]	Intent to Treat Population – Patients with genital erosion at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with SFQ Item 2 score of 0, compared between two groups of patients with presence/absence of genital fissure, erosion, and /or	Intent to Treat Population – Patients with genital fissure, erosion, and/or ulcer at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	ulcer[d]					
DLQI and psoriasis in perianal/ gluteal cleft area	Proportion of patients with DLQI (0,1), compared between two groups of patients with presence/absence of psoriasis in perianal/gluteal cleft area[d]	Intent to Treat Population – Patients with psoriasis in perianal and/or gluteal cleft area at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
DLQI and facial psoriasis	Proportion of patients with DLQI (0,1), compared between two groups of patients with presence/absence of facial psoriasis[d]	Intent to Treat Population – Patients with facial psoriasis at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
SF-36 MCS and genital fissure/ erosion/ ulcer	Proportion of patients with SF-36 MCS ≥2.5 improvement, compared between two groups of patients with presence/absence of genital fissure[d]	Intent to Treat Population – Patients with genital fissure at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
	Proportion of patients with SF-36 MCS ≥2.5 improvement, compared between	Intent to Treat Population – Patients with genital erosion at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

Measure	Variable	Population (Section 6.1.1)	Time Point	Analysis and Imputation Method (Sections 6.1.3 and 6.3)	Objective/ Endpoint Type	Analysis Type
	two groups of patients with presence/absence of genital erosion[d]					
	Proportion of patients with SF-36 MCS ≥2.5 improvement, compared between two groups of patients with presence/absence of genital fissure, erosion, and /or ulcer[d]	Intent to Treat Population – Patients with genital fissure, erosion, and/or ulcer at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
SF-36 MCS and psoriasis in perianal/ gluteal cleft area	Proportion of patients with SF-36 MCS ≥2.5 improvement, compared between two groups of patients with presence/absence of psoriasis in perianal/gluteal cleft area[d]	Intent to Treat Population – Patients with psoriasis in perianal and/or gluteal cleft area at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]
SF-36 MCS and facial psoriasis	Proportion of patients with SF-36 MCS ≥2.5 improvement, compared between two groups of patients with presence/absence of facial psoriasis[d]	Intent to Treat Population – Patients with facial psoriasis at baseline	Week 12	Fisher's exact test using NRI	Exploratory	Primary[e]

[a] Patients should complete the scales in the following order: 1. GPSS, 2. GPSIS, and 3. SFQ. An e-diary will be used to collect GPSS daily, GPSIS and SFQ weekly, up to Week 12 (Visit 7); thereafter, patients will answer these questions only at scheduled study visits. At Week 12 (Visit 7), patients will complete these scales on site before completion of any other assessments.

[b] The scores at each post-baseline visit for GPSS, GPSIS and SFQ, in the Blinded Treatment Period and Open-Label Treatment Period, are defined in Table.RHBQ.6.4.

[c] The scores at each calendar week in the Blinded Treatment Period, for GPSS, GPSIS and SFQ, are defined in Table.RHBQ.6.4.

[d] The variables are derived based on all patients in the analysis population, not for placebo or ixekizumab treated patients separately.

[e] The Primary Method for this analysis is the Fisher's exact test instead of the logistic regression model.

6.11. Analysis Methodology for Primary and Major Secondary Outcomes

The primary endpoint for this study is the proportion of patients achieving treatment success, defined as achieving sPGA of Genitalia (0,1) at Week 12 (end of the Blinded Treatment Period) and remaining on the initially randomized treatment (i.e. the effect attributable to the originally randomized treatment, ixekizumab Q2W. With this estimand any patient who discontinues treatment prior to Week 12 is considered a treatment failure through the use of the NRI. The numeric result for treatment success/failure is identical to NRI but interpreted differently and has different consequences for missing data. With NRI, as the acronym implies, there is explicit imputation of missing sPGA of Genitalia (0,1) outcomes. With treatment success/failure, discontinuation of study medication is considered a treatment failure because if patients cannot adhere to the medication they will not have sustained benefit from it. Therefore, every patient will have an observation for treatment success/failure and placebo, at the primary time point of Week 12 in all randomized patients) and there will be no missing data for this estimand and hence inferences will not depend on missing data assumptions. These attributes also apply to other endpoints involving use of NRI.

The major secondary endpoints for this study are

1. The proportion of patients achieving an overall sPGA (0,1) at Week 12.
2. The proportion of patients with at least a 3 point improvement in genital psoriasis itch NRS within the GPSS at Week 12. This will be calculated for patients who had baseline score of at least 3.
3. The proportion of patients achieved a SFQ Item #2 score of 0 or 1 at Week 12. This will be calculated for patients who had baseline score of at least 2.

The primary analysis method for the primary endpoint will be based on a logistic regression analysis with treatment and BSA category as factors in the model and compare the ixekizumab Q2W dosing versus placebo at Week 12 (Visit 7) using the ITT Population (Section 6.1.3).

The primary analysis method for the major secondary endpoint #1-3 will be based on a logistic regression analysis with treatment and BSA category as factors in the model and compare the ixekizumab Q2W dosing versus placebo at Week 12 (Visit 7) using the ITT population (Section 6.1.3).

The primary endpoint will be tested at 2-sided $\alpha = 0.05$. If the test for primary endpoint is significant, then the test for the major secondary endpoint #1 will be performed. If the test for major secondary endpoint #1 is significant, then the test for major secondary endpoint #2 will be performed. Similarly, the test for major secondary endpoint #3 will be performed only if all prior tests are significant. If a test is not significant, all subsequent tests will be considered not significant.

Other secondary and tertiary analyses will be performed on the primary and major secondary efficacy endpoints as detailed in Table RHBQ.6.5.

6.12. Other Secondary Efficacy Outcomes and Analysis methodology

There will be no adjustment for multiple comparisons for the other secondary efficacy outcomes. The other secondary endpoints are defined in Section 4.3. The analysis methods and corresponding populations are described in Table RHBQ.6.5.

6.13. Health Outcomes/Quality of Life Analyses

The health outcomes and quality of life measures for this study are:

- GPSS
- GPSIS
- SFQ
- Touch avoidance
- PatGA-Genital
- DLQI
- SF-36: MCS and PCS and individual domains
- CAPP genital sub-index

These measures are defined in Table RHBQ.6.4. The analysis methods and corresponding populations are described in Table RHBQ.6.5.

The psychometric analyses of the diary endpoints, including GPSS, GPSIS and SFQ, are outlined in Section 6.20.1.

6.14. Pharmacokinetic/Pharmacodynamic Methods

Details of the pharmacokinetic/pharmacodynamic (PK/PD) analyses can be found in a separate PK/PD analysis plan and are summarized briefly below.

Observed ixekizumab serum concentrations will be summarized by visits and corresponding time when sampling occurred.

As appropriate, the PK and the exposure-response relationship between ixekizumab exposure and clinically important efficacy measures (for example, sPGA or mGPASI) may be explored using graphical methods and/or a modeling approach. If a modeling approach is taken, data may be combined with data from other ixekizumab studies if appropriate.

The potential impact of immunogenicity on ixekizumab exposure and/or efficacy responses may be evaluated by graphical assessments, as appropriate, to compare drug exposure or efficacy responses between ADA negative and ADA positive patients at corresponding visits, or before and after ADA development for patients who developed ADA. Both treatment-emergent only and all ADA positive/negative patients may be explored. A similar approach may be taken if patients become neutralizing anti-drug antibody (NAb) positive.

Additional analyses may be performed upon receipt of the data. Data from this study may be combined with data from previous efficacy studies for additional population PK and/or exposure efficacy modelling if deemed appropriate.

6.15. Safety Analyses

Safety will be assessed by summarizing and analyzing AEs including adjudicated cerebro-cardiovascular events, QIDS-SR16, C-SSRS, laboratory analytes, and vital signs. The duration of exposure will also be summarized.

For the Blinded Treatment Period (Period 2), safety data will be summarized using the safety population. Treatment group comparisons between ixekizumab Q2W and placebo will be performed on categorical safety data using a Fisher's exact test and continuous safety data will be analyzed by an ANCOVA model with treatment group and baseline as factors.

For the Open-label Treatment Period (Period 3), safety data will be summarized according to the treatment group to which they were assigned in Period 2 and Period 3 (i.e. IXE80Q2W/IXE80Q4W and PBO/IXE80Q4W) and overall.

For the Post-Treatment Follow-up Period (Period 4), safety data will be summarized according to the treatment group in Period 2 or Period 3 (i.e. prior to entering Period 4) as described in Table RHBQ.6.1.

For safety analyses, the following baselines will be used:

- Treatment-emergent adverse events (TEAEs): baseline will be all results recorded during the baseline period (see Section 6.1.2 for definitions of the baseline period).
- Change from baseline to last observation and each scheduled post-baseline visit for laboratory and vital signs: baseline will be last non-missing assessment recorded during the baseline period (see Section 6.1.2 for definitions of the baseline period).
- Treatment-emergent abnormal laboratory and vital signs: baseline will be all results recorded during the baseline period (see Section 6.1.2 for definitions of the baseline period).
- Change from baseline to minimum or maximum: baseline will be all results recorded during the baseline period (see Section 6.1.2 for definitions of the baseline period).

6.15.1. Extent of Exposure

Duration of exposure to study drug during the Blinded Treatment Period (Period 2) will be summarized by treatment group for the safety population using descriptive statistics. Duration of exposure in the Open-label Treatment Period (Period 3) will be summarized by ixekizumab Q4W only and ixekizumab Q4W step up to Q2W groups and overall using the Open-label Treatment Population.

A by-patient listing of exposure duration with number of active injections and total dose will be provided.

Duration of Exposure to Study Treatment during the Blinded Treatment Period (Period 2)

The duration of exposure to ixekizumab Q2W in Period 2 will be calculated as:

$$Duration\ of\ exposure\ to\ Q2W\ (days) = Date\ of\ last\ visit\ (scheduled\ or\ unscheduled)\ in\ Period\ 2 - Date\ of\ first\ dose + 1$$

The number and percentage of patients in each of the following categories will be included in the summaries:

- >0, ≥7 days, ≥14 days, ≥30 days, ≥60 days, ≥90 days. Note that the same patients may be included in more than 1 category.
- >0 to <7 days, ≥7 to <14 days, ≥14 to <30 days, ≥30 to <60 days, ≥60 days to <90 days and ≥90 days

The summaries will also include the following information:

- Total exposure in patient years, calculated as:

$$Exposure\ in\ patient\ years = \frac{Sum\ of\ duration\ of\ exposure\ for\ Period\ 2\ (for\ all\ patients\ in\ treatment\ group)}{365.25}$$

- Number of active injections taken: this is derived using the response to the question "Was dose administered?" on the Exposure eCRF page and the actual dose description from IWRS study drug dispense dataset
- Total dose (in mg): this is calculated by the number of active injections taken during Period 2 multiplies by a dose of 80 mg

Duration of Exposure to Study Treatment during the Open-Label Treatment Period (Period 3)

The duration of exposure to ixekizumab Q4W only or ixekizumab Q4W step up to Q2W dosing in Period 3 will be calculated as:

$$Duration\ of\ exposure\ (days) = Date\ of\ last\ visit\ (scheduled\ or\ unscheduled)\ in\ Period\ 3] - Date\ of\ first\ Q4W\ dosing\ in\ Period\ 3 + 1$$

The number and percentage of patients in each of the following categories will be included in the summaries:

- >0, ≥7 days, ≥14 days, ≥30 days, ≥60 days, ≥90 days, ≥120 days, ≥180 days, and ≥280 days. Note that the same patients may be included in more than 1 category.
- >0 to <7 days, ≥7 to <14 days, ≥14 to <30 days, ≥30 to <60 days, ≥60 days to <90 days, ≥90 to <120 days, ≥120 to <180 days, ≥180 to <280 days, and ≥280 days.

Duration of Exposure to Ixekizumab Treatment during the Combined Treatment Period (Blinded Treatment Period (Period 2) and Open-Label Treatment Period (Period 3))

The duration of exposure to ixekizumab in Combined Period will be calculated, for patients who received at least one dose of ixekizumab, as:

$$Duration\ of\ exposure\ \ (days)$$
$$= Date\ of\ last\ visit\ (scheduled\ or\ unscheduled)\ in\ Combined\ Treatment\ Period]$$
$$-\ Date\ of\ first\ ixekizumab\ dose\ in\ Combined\ Treatment\ Period\ +\ 1$$

The number and percentage of patients in each of the following categories will be included in the summaries:

- >0, ≥7 days, ≥14 days, ≥30 days, ≥60 days, ≥90 days, ≥120 days, ≥180 days, ≥280 days ≥365 days, and ≥548 days. Note that the same patients may be included in more than 1 category.
- >0 to <7 days, ≥7 to <14 days, ≥14 to <30 days, ≥30 to <60 days, ≥60 days to <90 days, ≥90 to <120 days, ≥120 to <180 days, ≥180 to <280 days, ≥280 to <365 days, ≥365 to <548 days, and ≥548 days.

6.15.2. Adverse Events

Adverse events (AEs) will be classified based upon the latest version of the MedDRA. Adverse events will be recorded at every study visit. Any untoward condition starting on or after the date of informed consent will be considered an AE. Any pre-existing condition which worsens in severity on or after the date of informed consent will be considered and recorded as an AE on the AE eCRF page from the date of worsening onwards.

A treatment-emergent adverse event (TEAE) is defined as an AE that first occurred or worsened in severity after baseline and on or prior to the date of the last visit within the defined treatment period. Both the date/time of the event and the date/time of the dose (that is, injection) are considered when determining TEAEs. Treatment-emergent AEs will be assigned to the respective study periods in which they are considered treatment-emergent.

The MedDRA lowest level term (LLT) will be used when classifying AEs as treatment-emergent.

- The maximum severity recorded for each LLT prior to the first dose date/time in the treatment period will be used as the pre-treatment severity for that LLT. If an event during the baseline period has missing severity, and the event persists during the treatment period, then it will be considered as treatment-emergent, regardless of the post-baseline level of severity. Events with a missing severity during the treatment period will be considered treatment-emergent.
- Adverse events with a particular LLT will be classified as treatment-emergent if they first start on or after the first dose date/time in the treatment period (i.e., a patient has no pre-existing conditions with that lowest level term), or if the severity is greater than the pre-

treatment severity for that lowest level term. If a partial AE start date/time is present, the date/time will be compared as far as possible to the treatment start date/time in order to determine whether the event is treatment emergent or not. If there is any doubt, the event will be flagged as treatment emergent.

A follow-up emergent adverse event (FEAE) is defined as an event that first occurred or worsened in severity after the date of Visit 12 (that is, Week 52) or the ETV:

- The MedDRA LLT will be used when classifying AEs as follow-up emergent.
- For AEs that are ongoing at the date of Visit 12 or ETV, the maximum severity recorded for each LLT on or prior to the date of Visit 12 or ETV will be used as the follow-up baseline severity for that LLT.
- If a partial AE start date is present, the date will be compared as far as possible to the date of Visit 12 or ETV in order to determine whether the event is follow-up emergent or not. If there is any doubt, the event will be flagged as follow-up emergent, unless the same event was already counted as treatment-emergent during the treatment during a previous treatment period (Period 2 or Period 3).

Adverse events and TEAEs will be summarized and analyzed for the safety population for the Blinded Treatment Period (Period 2). The comparisons between treatment groups will be conducted using Fisher's exact test.

Adverse events and TEAEs will be summarized for the Open-label Treatment Population for the Open-label Treatment Period (Period 3).

The following will be presented for the Blinded and Open-label Treatment periods respectively:

- An overall summary of AEs including the number and percentage of patients who experienced TEAE, TEAE by maximum severity, death, SAE, TEAE related to study drug, discontinuations from the treatment due to an AE, and TEAEs of special interest
- TEAE by SOC and PT
- TEAE by PT
- TEAEs possibly related to study drug (evaluated by the investigator), by SOC and PT
- TEAE by maximum severity, SOC, and PT.

In addition, the overall summary of AEs will also be presented for the patients who received at least one dose of ixekizumab, during the Combined Treatment Period (Blinded Treatment Period [Period 2] and Open-label Treatment period [Period 3]). The summary table includes the number and percentage of patients who experienced TEAE, TEAE by maximum severity, death, SAE, TEAE related to study drug, discontinuations from the treatment due to an AE, and TEAEs of special interest, during the Combined Treatment Period.

Follow-up emergent adverse events will be summarized for the follow-up population for the Post-Treatment Follow-up Period (Period 4):

- FEAE by PT

In general, for all AE related summaries, the number and percentage of patients experiencing the events will be presented by treatment group. The events will be ordered by decreasing frequency in the total ixekizumab group, followed by Q2W and placebo in the Blinded Treatment Period and for total ixekizumab group in the Open-label Treatment Period within SOC and/or PT for sorting. For events that are gender-specific (as defined by MedDRA), the denominator and computation of the percentage will include only patients from the given gender.

A by-patient listing of all AEs will be provided.

6.15.2.1. Common Adverse Events

Common TEAEs are those TEAEs that occurred in ≥1% before rounding of total ixekizumab treated patients.

The following summaries (including treatment group comparison) will be provided for common TEAEs based on the safety population for Period 2 and based on Open-label Treatment Population for Period 3:

- Common TEAEs by SOC and PT
- Common TEAEs by PT
- Common TEAEs by maximum severity, SOC, and PT.

6.15.3. Deaths, Other Serious Adverse Events, and Other Notable Adverse Events

By-patient listings of deaths, serious AEs (SAEs), and AEs leading to discontinuation will be provided, respectively. All deaths will be included, regardless of the investigator's or the sponsor's judgment about causality, including:

- any deaths occurring during participation in the study in the database for which data are being presented
- any deaths occurring after a patient leaves (is discontinued from or completed) the study in the database for which data are being presented if:
    - the death is the result of a process initiated during the study, regardless of when it actually occurred, or
    - the death occurs during the Period 4 after discontinuation of study drug.

An SAE is any AE that results in one of the following outcomes: death, life-threatening, initial or prolonged hospitalization, disability or permanent damage, congenital anomaly or birth defect, or any other serious/important medical events.

The following summary tables (including treatment group comparison) will be provided for the safety population for Period 2 and the Open-label Treatment Population for Period 3:

- SAEs by PT

- AEs that lead to treatment discontinuation (including death) by PT

A follow-up emergent serious adverse event (FESAE) is defined as an SAE that first occurred or worsened in severity after the date of Visit 12 (that is, Week 52) or the ETV. The following summary tables will be provided for the follow-up population for the Post-Treatment Follow-up Period (Period 4):

- FESAE by PT
- FEAEs that lead to treatment discontinuation (including death) by PT.

6.15.3.1. Special Safety Topics including Adverse Events of Special Interest

Safety information on special topics including AEs of special interest (AESI) will be presented by treatment group and by study period. Table RHBQ.6.6 provides the definitions/derivations and analyses methods (including analyses, summaries and by-patient listings) of special safety topics including AESIs.

Potential AESIs will be identified by a standardized MedDRA query (SMQ) or a Lilly-defined MedDRA PT listing. Preferred terms within an SMQ will be classified as broad and narrow. In the Lilly-defined MedDRA PT listings, Lilly has provided the broad and narrow classifications. The Lilly-defined broad terms are for a more sensitive search of potential events of interest and the Lilly-defined narrow terms are for a more specific search. Therefore, the summaries will include the classifications of broad term (same as pooling narrow and broad terms together) and narrow term.

In the event that the listing of terms or analyses changes for a special safety topic, it will be documented in the IXE PSAP which will supersede this document; it will not warrant an amendment to the individual study SAP.

Fisher's exact tests will be used to compare the treatment groups for the safety population during the Blinded Treatment Period (Period 2).

For the Open-label Treatment Period (Period 3) and the Post-Treatment Follow-up Period (Period 4), summaries will be provided using the Open-label Treatment Population and the Post-Treatment Follow-up Population, respectively.

Table RHBQ.6.6. Definitions and Analyses of Special Safety Topics including Adverse Events of Special Interest

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
Hepatic	Hepatic AE analysis will include events that are potentially drug-related hepatic disorders by using the Medical Dictionary for Regulatory Activities (MedDRA) PTs contained in any of the following standardized MedDRA query (SMQ) or sub-SMQ as defined in MedDRA: • Broad and narrow terms in the Liver related investigations, signs and symptoms (20000008) • Broad and narrow terms in the Cholestasis and jaundice of hepatic origin (20000009) • Broad and narrow terms in the Hepatitis, non-infectious (20000010) • Broad and narrow terms in the Hepatic failure, fibrosis and cirrhosis and other liver damage (20000013) • Narrow terms in the Liver-related coagulation and bleeding disturbances (20000015)	**Period 2 (Fisher's exact test)**: TEAE by PT within SMQ or sub-SMQ, SAE by PT within SMQ or sub-SMQ, AE leading to treatment discontinuation by PT within SMQ or sub-SMQ **Period 3 (Summary)**: TEAE by PT within SMQ or sub-SMQ, SAE by PT within SMQ or sub-SMQ, AE leading to treatment discontinuation by PT within SMQ or sub-SMQ **Period 4 (Summary)** FEAE by PT within SMQ or sub-SMQ **Listing:** TEAE
	Elevations in hepatic laboratory tests (ALT, AST, ALP, total bilirubin) using performing lab reference ranges are defined as: • Include scheduled visits, unscheduled visits, and repeat measurements. • Alanine aminotransferase (ALT) or aspartate aminotransferase (AST): maximum post-baseline measurement ≥3 times (3×), 5 times (5×), 10 times (10×), and 20 times (20×) the performing lab upper limit of normal (ULN) for all patients with a post-baseline value. o The analysis of 3× ULN will contain 4 subsets: patients whose non-missing maximum baseline value is ≤1× ULN, >1× ULN to <3× ULN, ≥3× ULN, or missing. o The analysis of 5× ULN will contain 5 subsets: patients whose non-missing maximum baseline value is ≤1× ULN, >1× ULN to <3× ULN, ≥3× ULN to <5× ULN, ≥5× ULN, or missing. o The analysis of 10× ULN will contain 6 subsets: patients whose non-missing maximum baseline value is ≤1× ULN, >1×ULN to <3× ULN, ≥3× ULN to <5×	**Period 2 (Fisher's exact test)**: Elevations in hepatic laboratory tests: maximum baseline category to abnormal maximum post-baseline category **Period 3 (Summary)**: Elevations in hepatic laboratory tests: maximum baseline category to abnormal maximum post-baseline category **Period 4 (Summary)**: Elevations in hepatic laboratory tests: maximum baseline category to abnormal maximum post-baseline category

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	ULN, ≥5× ULN to <10× ULN, ≥10× ULN, or missing. o The analysis of 20× ULN will contain 7 subsets: patients whose non-missing maximum baseline value is ≤1×ULN, >1×ULN to <3× ULN, ≥3× ULN to <5× ULN, ≥5× ULN to <10× ULN, ≥10× ULN to <20× ULN, ≥20× ULN, or missing. • The number and percentages of patients with a total bilirubin measurement ≥1.5 times (1.5×), and ≥2 times (2×) the performing lab ULN during the treatment period will be summarized for all patients with a post-baseline value. o The analysis of 1.5× ULN will contain four subsets: patients whose non-missing maximum baseline value is ≤1× ULN, patients whose maximum baseline is >1× ULN but <1.5× ULN, patients whose maximum baseline value is ≥1.5× ULN, and patients whose baseline values are missing. o The analysis of 2× ULN will contain five subsets: patients whose non-missing maximum baseline value is ≤1× ULN, patients whose maximum baseline is >1× ULN but <1.5× ULN, patients whose maximum baseline is ≥1.5× ULN but <2× ULN, patients whose maximum baseline value is ≥2× ULN, and patients whose baseline values are missing. • ALP: maximum post-baseline measurement >1.5 times (1.5×) the performing lab ULN for all patients with a post-baseline value, and divided into 4 subsets: patients whose non-missing maximum baseline value is ≤1× ULN, >1× ULN to ≤1.5× ULN, >1.5× ULN, or missing.	
	Shift for ALT, AST, and total bilirubin from maximum baseline to maximum post-baseline will be produced with the requirements using performing lab reference ranges: • Include scheduled visits, unscheduled visits, and repeat measurements. • Use the maximum non-missing value in the baseline period. • Use the maximum non-missing post-baseline value within each study period. • Categories are: o ALT: ≤1× ULN, >1 to <3× ULN, ≥3 to <5× ULN, ≥5 to <10× ULN, ≥10 to <20× ULN, and ≥20× ULN o AST: ≤1× ULN, >1 to <3× ULN, ≥3 to <5× ULN, ≥5 to <10× ULN, ≥10× to <20× ULN and ≥20× ULN o Total bilirubin: ≤1× ULN, >1 to <1.5× ULN, ≥1.5 to <2× ULN, ≥2× ULN o ALP; ≤1× ULN, >1 to ≤1.5× ULN, >1.5× ULN. • With additional categories:	**Period 2 (Summary)**: Shifts from maximum baseline to maximum post-baseline category **Period 3 (Summary)**: Shifts from maximum baseline to maximum post-baseline category

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	○ Decreased: post-baseline category < baseline category ○ Increased: post-baseline category > baseline category ○ Same: post-baseline category = baseline category	
	Elevated hepatic criteria: maximum ALT ≥3× ULN and maximum total bilirubin ≥2× ULN, using performing lab reference ranges. Listing of patients who meet any of the following criteria: • Elevated hepatic criteria: defined as maximum ALT ≥3× ULN, and maximum total bilirubin ≥ 2× ULN • An ALT or AST ≥ 3× ULN • An alkaline phosphatase (ALP) ≥ 2× ULN • A total bilirubin ≥ 2× ULN The listing will include: patient demographics, concomitant medications, ALT/AST/ALP/total bilirubin/GGT by visit, treatment start and stop dates, and reason for treatment discontinuation	**Period 2 (Fisher's exact test)**: Elevated hepatic criteria **Period 3 (Summary)**: Elevated hepatic criteria **Period 4 (Summary)**: Elevated hepatic criteria **Listing**: Elevations in hepatic laboratory tests
	Evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plot: use maximum ALT measurement and maximum total bilirubin measurement with patients having at least one post-baseline ALT and total bilirubin, which contributes one point to the plot. The measurements do not need to be taken at the same blood draw.	**Period 2**: eDISH plot **Period 3:** eDISH plot **Period 4:** eDISH plot
Cytopenias	Cytopenias are defined using the PTs from the following 2 sub-SMQs of the Haematopoietic cytopenias SMQ (20000027) as specified in MedDRA: • Broad and narrow terms in the Haematopoietic leukopenia (20000030) • Broad and narrow terms in the Haematopoietic thrombocytopenia (20000031)	**Period 2 (Fisher's exact test)**: TEAE by PT within sub-SMQ, SAE by PT within sub-SMQ, AE leading to treatment discontinuation by PT within sub-SMQ **Period 3 (Summary)** TEAE by PT within sub-SMQ, SAE by PT within sub-SMQ, AE leading to treatment discontinuation by PT within sub-SMQ **Period 4 (Summary)**: FEAE by PT within sub-SMQ

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
		Listing: TEAE
Infections	Infections are events including infections, serious infections, opportunistic infections, infections that require therapeutic intervention (antibiotics, antivirals, antifungals, etc.), and any events involving reactivation of TB or hepatitis B or hepatitis C. Specifically, infections are defined using all the PTs from the Infections and Infestations System Organ Class (SOC) as specified in MedDRA.	**Period 2 (Fisher's exact test)**: TEAE by PT, TEAE by maximum severity by PT, SAE by PT, AE leading to treatment discontinuation by PT **Period 3 (Summary)**: TEAE by PT, TEAE by maximum severity by PT, SAE by PT, AE leading to treatment discontinuation by PT
	Anti-infective medications are defined in Appendix 5 of Ixekizumab Program Safety Analysis Plan Version 7 (IXE PSAP V7) including antibiotics, antifungals, antivirals, or antiprotozoals. Listing of patients experiencing a TEAE of infections will be provided including the following additional information: anti-infective medications use (if treated) with medication start/end dates, indication for use, and route; minimum post-baseline value within treatment Period 2 for leukocytes, platelets, lymphocytes, and absolute neutrophils.	**Listing:** TEAE with anti-infective medications.
	The opportunistic infections (OI) are defined as: • The Lilly specified list is contained in Appendix 11 of the IXE PSAP V7. The narrow terms are considered opportunistic infections. Medical review of broad terms is needed for final determination of patients with OIs. Listing of patients experiencing a TEAE of OIs will be provided including the following additional information: source of identification (CRF or Lilly specified list), primary/secondary site of infection, primary/secondary infection type, primary/secondary identified by a laboratory diagnostic test (Yes/No), acquired in a Health care setting (Yes/No).	**Listing:** TEAE of OIs

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	The duration of each common (≥1% of total ixekizumab) TEAE PT of Infections and Opportunistic Infections is defined as: Duration of treatment-emergent AE Infections (in weeks) = (End date of AE – Start date of AE + 1) / 7. Only TEAEs of infections beginning during treatment Period 2 will be included in the summary. If an AE has not ended by the date of completion of the treatment Period 2, or date of early discontinuation, it will be censored as of that date (last visit within the treatment Period 2, or date of early discontinuation). If a patient has multiple episodes of the same TEAE, the episode with the greatest severity will be used for the duration of event calculation. If a patient has multiple episodes of the same TEAE with the same severity, the episode with the longest duration will be used for the duration of event calculation.	**Period 2 (Summary)**: Duration of Common TEAE – Infections Duration of Common TEAE – Opportunistic Infections
Allergic Reactions / Hypersensitivities	Allergic reactions/hypersensitivity events will be categorized as either anaphylaxis or non-anaphylaxis events (these will refer to events that are not localized to the site of injection) and summarized separately. Medical reviews are needed for final determination of patients with allergic reactions/hypersensitivities. Allergic Reactions/Hypersensitivity Events, Anaphylaxis: Anaphylaxis has been broadly defined as "a serious allergic reaction that is rapid in onset and may cause death" (Sampson et al. 2006). Identification of cases of potential anaphylaxis from the clinical trial data involves two screening criteria, one designed to specifically identify cases (following Criterion 1) based on narrow terms from the MedDRA SMQ for anaphylactic reaction (20000021), and the second to identify possible cases, following Criterion 2 as defined by Sampson et al. (2006). 1) Criterion 1 for anaphylaxis is defined by the presence of a TEAE based on the following MedDRA PTs from the anaphylactic reaction SMQ: • Anaphylactic reaction • Anaphylactic shock • Anaphylactoid reaction • Anaphylactoid shock • Kounis Syndrome • Type 1 hypersensitivity 2) Criterion 2 for anaphylaxis requires having TEAEs from two or more of four categories	**Period 2 (Fisher's exact test)**: TEAE by PT within Category, TEAE by maximum severity by PT within Category, SAE by PT within Category, AE leading to treatment discontinuation by PT within Category, **Period 3 (Summary)**: TEAE by PT within Category, TEAE by maximum severity by PT within Category, SAE by PT within Category, AE leading to treatment discontinuation by PT within Category,

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	of AEs as described by Sampson et al. (2006). Occurrence of these events should be nearly coincident; based on recording of events on CRFs. All qualifying events must be within 1 day of study drug injection. The 4 categories to be considered in Criterion 2 are: • Category A: Involvement of the skin-mucosal tissue • Category B: Respiratory compromise • Category C: Reduced blood pressure or associated symptoms • Category D: Persistent gastrointestinal symptoms The specific MedDRA PTs covered by each of these Criterion 2 categories are shown in Appendix 7 of IXE PSAP V7. Summaries of Criterion 2 anaphylactic TEAEs will be provided by the specific combination of categories as follows: • AB: events based on meeting Category A and Category B (but no other category) • AC: events based on meeting Category A and Category C (but no other category) • AD: events based on meeting Category A and Category D (but no other category) • BC: events based on meeting Category B and Category C (but no other category) • BD: events based on meeting Category B and Category D (but no other category) • CD: events based on meeting Category C and Category D (but no other category) • ABC: events based on meeting Category A, Category B and Category C (but no other category) • ABD: events based on meeting Category A, Category B and Category D (but no other category) • ACD: events based on meeting Category A, Category C and Category D (but no other category) • BCD: events based on meeting Category B, Category C and Category D (but no other category) • ABCD: events based on meeting each of the 4 Criterion 2 categories Summaries of treatment-emergent anaphylactic AEs will be provided for patients meeting each of the 2 criteria and for patients who meet either criteria overall. Severity of treatment-emergent Criterion 2 anaphylactic AEs will be based on the maximum severity of the specific events met by the patient. Maximum severity of an (or overall) treatment-emergent	

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	anaphylactic AE will be based on the maximum severity within Criterion 1 and/or Criterion 2. Allergic Reactions/Hypersensitivity Events, Non-Anaphylaxis: TEAEs of allergic reaction/hypersensitivity categorized as non-anaphylaxis events are defined by the narrow terms within Hypersensitivity SMQ (20000214) excluding the PTs noted in Appendix 8 of the IXE PSAP V7 and excluding the anaphylactic events as defined above.	
	A by-patient listing will be provided for all patients experiencing TEAE of allergic reactions/hypersensitivities at any time, including status/criterion of anaphylaxis or non-anaphylaxis, and the associated information collected on *Allergic / Hypersensitivity Reaction Follow-Up* eCRF page if identified by the investigator.	**Listing:** TEAE including information collected on *Allergic / Hypersensitivity Reaction Follow-Up* eCRF page
Injection Site Reactions	Injection site reaction is defined using the PTs from the MedDRA HLT of Injection site reactions as specified by MedDRA excluding the following 10 PTs: 1) Embolia cutis medicamentosa 2) Injection site joint discomfort 3) Injection site joint effusion 4) Injection site joint erythema 5) Injection site joint infection 6) Injection site joint inflammation 7) Injection site joint movement impairment 8) Injection site joint pain 9) Injection site joint swelling 10) Injection site joint warmth If the AESI trigger CRF form question "Since the last visit did an Injection Site reaction occur?" is answered "Yes", the Injection Site Reaction follow-up form will be completed. Some additional information in this follow-up form is described in Appendix 4 of the IXE PSAP V7.	**Period 2 (Fisher's exact test)**: TEAE by PT within HLT, TEAE by maximum severity by PT within HLT, SAE by PT within HLT, AE leading to treatment discontinuation by PT within HLT TEAE identified by the investigator PT within HLT: by maximum severity, by maximum redness category, by maximum swelling category, by maximum pain category **Period 3 (Summary)**: TEAE by PT within HLT, TEAE by maximum severity by PT within HLT, SAE by PT within HLT, AE leading to treatment discontinuation by PT within HLT

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
		TEAE identified by the investigator PT within HLT: by maximum severity, by maximum redness category, by maximum swelling category, by maximum pain category **Listing:** TEAE including information collected on *Injection Site Reaction* eCRF page
Cerebro-cardiovascular Events	Cerebro-cardiovascular events will be externally adjudicated by the Clinical Events Committee (CEC) at the Cleveland Clinic, as outlined in the Manual of Operations. Investigator-reported events will be selected for adjudication, based on the assigned MedDRA PT and the criteria outlined in the listing provided in Appendix 9 of the IXE PSAP V7. The CEC will adjudicate investigator-reported events selected for adjudication and render an assessment as to whether the event represents a confirmed event (meeting the event definition with all necessary documentation), a non-event (does not meet the event definition and likely represents an alternative or nonevent diagnosis), or lacks sufficient documentation for confirmation of an event. All events which qualify for CEC adjudication will be used for the analysis of cerebro-cardiovascular events. The categories of adjudicated events used for the analysis will include the following: • Cardiovascular o Death (Cardiovascular) o Myocardial Infarction (MI) o Hospitalization for Unstable Angina o Hospitalization for Heart Failure o Serious Arrhythmia o Hospitalization for Hypertension	**Period 2 (Fisher's exact test)**: TEAE by PT within Subcategory, SAE by PT within Subcategory, AE leading to treatment discontinuation by PT within Subcategory **Period 3 (Summary)**: TEAE by PT within Subcategory, SAE by PT within Subcategory, AE leading to treatment discontinuation by PT within Subcategory **Listing:** TEAE

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	○ Resuscitated Sudden Death ○ Cardiogenic Shock due to Myocardial Infarction ○ Coronary Revascularization Procedure • Neurologic ○ Cerebrovascular Event: Transient Ischemic Attack or Stroke (Hemorrhagic, Ischemic and Undetermined) • Peripheral Vascular Events ○ Peripheral Arterial Event ○ Peripheral Revascularization Procedure Events will be analyzed using MedDRA PT nested within the CEC assessment (confirmed event, no event, or insufficient documentation for event determination) and the subcategory. Subtypes of stroke (Hemorrhagic Stroke, Ischemic Stroke, and Undetermined Stroke Type) will be displayed in the analyses nested within Cerebrovascular Event. Subtypes of Serious Arrhythmia (Atrial Arrhythmia, Ventricular Arrhythmia, Heart Block, Other, Unknown) will be displayed nested within Serious Arrhythmia.	
Major Adverse Cardiovascular Events (MACE)	Major Adverse Cerebro-Cardiovascular Events (MACE) in general are a subset of the cerebro-cardiovascular events (requiring adjudication as defined above). MACE events are defined as follows (see Appendix 9 of the IXE PSAP V7): • Vascular Death (including cardiovascular and cerebro-vascular causes excluding hemorrhagic deaths outside of the central nervous system) • Non-fatal myocardial infarction • Non-fatal stroke (subtypes: hemorrhagic stroke, ischemic stroke, undetermined stroke type)	**Period 2 (Fisher's exact test)**: TEAE by maximum severity by PT within Category **Period 3 (Summary)**: TEAE by maximum severity by PT within Category **Listing:** TEAE
Malignancies	Malignancy is defined using PTs from the Malignant or unspecified tumors SMQ as specified in MedDRA (SMQ: 20000091, which includes the sub-SMQs: 20000195 [Tumours of unspecified malignancy] and 20000194 [Malignant tumours]). Events will be summarized by the following categories: • Non-Melanoma Skin Cancer (NMSC) ○ Basal Cell Carcinoma, PTs include:	**Period 2 (Fisher's exact test)**: TEAE by PT within Category, SAE by PT within Category, AE leading to treatment discontinuation by PT within Category

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	▪ Basal cell carcinoma ▪ Basosquamous carcinoma ▪ Basosquamous carcinoma of skin o Squamous Cell Carcinoma, PTs include: ▪ Squamous cell carcinoma of skin ▪ Bowen's disease ▪ Lip squamous cell carcinoma ▪ Skin squamous cell carcinoma metastatic ▪ Keratoacanthoma • Malignancies excluding NMSC: all PTs in the Malignant or unspecified tumors SMQ excluding the 8 defined NMSC PTs.	**Period 3 (Summary)**: TEAE by PT within Category, SAE by PT within Category, AE leading to treatment discontinuation by PT within Category **Listing:** TEAE
Depressions and Suicide/Self-injury	Depression is defined using the PTs from the Depression and suicide/self-injury SMQ as specified in MedDRA (SMQ: 20000035, which includes the sub-SMQs: 20000037 [Suicide/self-injury] and 20000167 [Depression (excl suicide and self- injury)]).	**Period 2 (Fisher's exact test)**: TEAE by PT within SMQ and sub-SMQ, SAE by PT within SMQ and sub-SMQ, AE leading to treatment discontinuation by PT within SMQ and sub-SMQ **Period 3 (Summary)**: TEAE by PT within SMQ and sub-SMQ, SAE by PT within SMQ and sub-SMQ, AE leading to treatment discontinuation by PT within SMQ and sub-SMQ **Listing:** TEAE
Interstitial Lung Disease (ILD)	ILD is defined using the following terms: • Broad and narrow terms in the Interstitial lung disease SMQ (20000042) • Additional 6 PTs from Eosinophilic pneumonia SMQ (20000157): o Angiolymphoid hyperplasia with eosinophilia (Narrow)	**Listing**: TEAE

Special Safety Topic	Definition / Derivation	Analysis / Summary / Listing
	○ Eosinophilic bronchitis (Narrow) ○ Hypereosinophilic syndrome (Narrow) ○ Loeffler's syndrome (Narrow) ○ Pulmonary eosinophilia (Narrow) ○ Pulmonary vasculitis (Narrow)	
Inflammatory Bowel Disease (IBD)	Inflammatory Bowel Disease (IBD) will be identified using the following subcategory and MedDRA PTs. The narrow terms are considered IBD. Medical reviews of patients identified with broad terms are needed for final determination of patients with IBD. IBD (Narrow terms) 1. Inflammatory Bowel Disease: Inflammatory bowel disease 2. Crohn's Disease: Crohn's disease 3. Ulcerative Colitis: Acute haemorrhagic ulcerative colitis; Colitis ulcerative; Proctitis ulcerative Non-Specific Terms: The PTs in this category are listed in Appendix 12 of the IXE PSAP V7	**Listing**: TEAE

Abbreviations: AE = adverse event; AESI = adverse events of special interest; ALP = alkaline phosphatase; CRF = case report form; eCRF = electronic case report form; FEAE = follow-up emergent adverse event; PT = preferred term; SAE = serious adverse event; TB = tuberculosis; TEAE = treatment emergent adverse event.

6.15.4. Clinical Laboratory Evaluation

Clinical laboratory assessments include hematology, serum chemistry, urinalysis, and safety related immune markers such as neutrophil counts. Continuous laboratory tests will be summarized as changes from baseline to last observation for patients who have both baseline and at least one post-baseline result for Period 2, Period 3, and Period 4, respectively:

- The scheduled visits/measurements will be included. The unscheduled visits and the repeated measurements taken at the same visit will be excluded.
- Both international system of unit (SI) and conventional unit will be summarized when different.
- For the safety population in the Blinded Treatment Period (Period 2), the comparisons between treatment groups will be conducted using an ANCOVA method with treatment group and baseline value in the model.
- Data will be analyzed based on original scale.

Laboratory test observed values at each visit (starting at baseline) and change from baseline to each scheduled visit, respectively, will be displayed in box plots for patients who have both a baseline and at least one post-baseline result. These box plots will be used to evaluate trends over time and to assess a potential impact of outliers on central tendency summaries.

The scheduled visits/measurements will be included. The unscheduled visits and the repeated measurements taken at the same visit will be excluded.

- The displays with both SI and conventional units will be provided when different.
- The following summary statistics will be included as a table below the box plot: number of patients with a baseline and at least one post-baseline result, mean, standard deviation, minimum, Q1, median, Q3, and maximum.
- Data will be summarized based on original scale.
- On the box plots of the laboratory test observed values, the lines of the reference ranges/limits (by using the large clinical trial population based reference limits, that is, Lilly reference ranges) will be added. In cases where limits vary across age and gender, the lowest of the high limits and the highest of the low limits will be used.
- The number and percentage of patients with a treatment-emergent or follow-up emergent abnormal, high, or low for laboratory tests will be summarized by treatment group for each study period. The comparisons between treatment groups will be conducted using Fisher's exact test for the safety population in the Blinded Treatment Period (Period 2). All scheduled, unscheduled and repeated measurements taken at the same visit will be included.
- In general, large clinical trial population based reference limits (that is, Lilly reference ranges) will be used to define the low and high limits since it is generally desirable for

limits used for analyses to have greater specificity (identify fewer false positive cases) than reference limits used for individual subject management. In the case when the reference limits based on the large clinical trial population is not available for a laboratory measure, performing lab reference ranges will be used. Alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin, and alkaline phosphatase (ALP), neutrophils, leukocytes, platelets and lymphocytes, will not be included in the treatment-emergent abnormal, high, or low summary as a separate analysis addressing the risk of liver injury is described in Section 6.15.3.1 in which performing lab reference ranges are used.

- Note that the ranges are defined by a lower limit of normal (LLN) and an upper limit of normal (ULN). A result that is greater than or equal to the LLN and less than or equal to the ULN is considered to be within the normal ranges.
- For categorical laboratory tests:
    - Treatment-emergent abnormal value is defined as a change from normal at all baseline visits to abnormal at any time post-baseline during the Blinded Treatment Period.
    - Treatment-emergent abnormal value is defined as a change from normal at baseline to abnormal at any time post-baseline during the Open-label Treatment Period.
    - Follow-up emergent abnormal result is defined as a change from normal at baseline to abnormal at any time during the follow-up period.
- For continuous laboratory tests:
    - Treatment-emergent high value is defined as a change from a value less than or equal to the ULN at all baseline visits to a value greater than the ULN at any time post-baseline during the Blinded Treatment Period.
    - Treatment-emergent low value is defined as a change from a value greater than or equal to the LLN at all baseline visits to a value less than the LLN at any time post-baseline during the Blinded Treatment Period.
    - Treatment-emergent high value is defined as a change from a value less than or equal to the ULN at baseline to a value greater than the ULN at any time post-baseline during the Open-label Treatment Period.
    - Treatment-emergent low value is defined as a change from a value greater than or equal to the LLN at baseline to a value less than the LLN at any time post-baseline during the Open-label Treatment Period.
    - Follow-up emergent high value is defined as a change from a value less than or equal to the ULN at baseline to a value greater than the ULN at any time post-baseline during the follow-up period.

- Follow-up emergent low value is defined as a change from a value greater than or equal to the LLN at baseline to a value less than the LLN at any time post-baseline during the follow-up period.

A by-patient listing of laboratory test values will be provided. A listing of laboratory tests reference ranges (both large clinical trial population based reference limits and performing lab reference ranges) will be provided. By-patient listing of abnormal laboratory test results (criteria defined in the shift tables excluding the normal category) for parameters of special interest (hepatic, leukocytes and platelets) will be provided.

6.15.4.1. Leukocytes (WBC) and Platelets

Further analyses will be conducted for total leukocytes, neutrophils, platelets, lymphocytes, monocytes, eosinophils, and basophils. Neutrophils will be defined as absolute neutrophils (derived by adding segmented neutrophils and band neutrophil). Shift table will be produced showing the number and percentage of patients shifting from baseline to a minimum post-baseline result in each relevant category by treatment group for Period 2, Period 3 and Period 4, respectively.

- Scheduled visits, unscheduled visits, and repeated measurements taken at the same visit will be included.
- Baseline is defined as the minimum result during the defined baseline period or baseline.
- Use the minimum non-missing post-baseline value within each study period.
- The parameters and categories are:
  - Leukocytes; $\geq$1×LLN, <1×LLN to $\geq 3.0\times 10^9$/L, $<3.0\times 10^9$/L to $\geq 2.0\times 10^9$/L, $<2.0\times 10^9$/L to $\geq 1.0\times 10^9$/L, and $<1.0\times 10^9$/L
  - Neutrophils (absolute neutrophils); $\geq$1×LLN, <1×LLN to $\geq 1.5\times 10^9$/L, $<1.5\times 10^9$/L to $\geq 1.0\times 10^9$/L, $<1.0\times 10^9$/L to $\geq 0.5\times 10^9$/L, and $<0.5\times 10^9$/L
  - Platelets; $\geq$1×LLN, <1×LLN to $\geq 75.0\times 10^9$/L, $<75.0\times 10^9$/L to $\geq 50.0\times 10^9$/L, $<50.0\times 10^9$/L to $\geq 25.0\times 10^9$/L, and $<25.0\times 10^9$/L
  - Lymphocytes; $\geq$1×LLN, < 1×LLN to $\geq 0.8\times 10^9$/L, $<0.8\times 10^9$/L to $\geq 0.5\times 10^9$/L, $<0.5\times 10^9$/L to $\geq 0.2\times 10^9$/L, and $<0.2\times 10^9$/L
- The above LLNs are defined as:
  - Leukocytes: LLN = 4.0×10^9/L
  - Neutrophils: LLN = 2.0×10^9/L
  - Lymphocytes: LLN = 1.1×10^9/L
  - Platelets: LLN = 150×10^9/L
- With additional categories:

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

The change from minimum baseline to minimum post-baseline result for each of these leukocytes and platelets will be summarized graphically using a box plot for Period 2, Period 3, and Period 4, respectively.

6.15.4.2. Neutrophil Follow-up

Neutrophil counts will be followed throughout the study. Patients will continue in Period 4 until their neutrophil counts have recovered.

The neutrophil follow-up analysis will be conducted on the neutrophil follow-up population defined as patients who have an absolute neutrophil count <1500 cells/μL (SI units: $<1.5\times 10^9$/L) at the last scheduled visit or early termination visit prior to entering the Post-Treatment Follow-up Period (Period 4) and less than the patient's baseline absolute neutrophil count (that is, prior to first injection at Week 0). These patients are monitored during the Period 4 until neutrophil recovery.

Neutrophil clinical recovery is defined as an absolute neutrophil count ≥1500 cells/μL (SI units: $\geq 1.5\times 10^9$/L) or greater than or equal to a patient's minimum absolute neutrophil count prior to first study drug injection at Week 0.

If a patient's neutrophil count has not recovered, within 12 weeks after entering the follow-up period (Visit 802), the patient will return for Visit 803 (12 weeks after Visit 802). Additional visits may be required for appropriate patient management depending upon the degree of neutropenia. If at Visit 802, a patient's has met the criteria for neutrophil recovery, the patient's participation in the study will be considered complete unless the investigator deems additional follow-up may be necessary.

The number and percentage of patients achieving neutrophil clinical recovery will be presented by treatment group and week interval for neutrophil follow-up population for Post-Treatment Follow-up Period (Period 4). The number and percentage of patients with an absolute neutrophil cell count that is at least 25%, 50%, 75%, or 100% of the patient's baseline absolute neutrophil count (that is, prior to first injection at Week 0), irrespective of absolute neutrophil minimum, will be included in the summary.

6.15.5. Vital Signs and Other Physical Findings

Analyses will be conducted on vital signs and physical characteristics including systolic blood pressure (mm Hg), diastolic blood pressure (mm Hg), pulse (bpm), weight (kg), body mass index (BMI) (kg/m^2), and waist circumference (cm). By-patient listings of vital signs and physical characteristics will be provided.

Change from baseline to last observation for vital signs and physical characteristics will be summarized for patients who have both baseline and at least one post-baseline result, for Period 2, Period 3 and Period 4, respectively:

The scheduled visits/measurements will be included. The unscheduled visits and the repeated measurements taken at the same visit will be excluded.

- For the safety population in the Blinded Treatment Period (Period 2), the comparisons between treatment groups will be conducted using an ANCOVA with treatment group and baseline value in the model.
- Data will be analyzed based on original scale.

For vital signs and physical characteristics, the observed values at each visit (starting at baseline) and change from baseline to each scheduled visit, respectively, will be displayed in box plots for patients who have both a baseline and at least one post-baseline result. These box plots will be used to evaluate trends over time and to assess a potential impact of outliers on central tendency summaries.

- The scheduled visits/measurements will be included. The unscheduled visits and the repeated measurements taken at the same visit will be excluded.
- The following summary statistics will be included as a table below the box plot: number of patients with a baseline and at least one post-baseline result, mean, standard deviation, minimum, Q1, median, Q3, and maximum.
- Data will be summarized based on original scale.

To assess the effect of administration of study drug on vital signs (blood pressures and pulse rate) among patients, at Week 0, vital signs will be measured before the first injection and 1 hour after the injection. The box plots will be produced for pre-dose and post-dose vital signs at Week 0 (Visit 2) by treatment group for safety population.

The number and percentage of patients with treatment-emergent or follow-up emergent high or low vital sign and weight at any time for Period 2, Period 3 and Period 4, respectively, will be summarized. The comparisons between treatment groups will be conducted using Fisher's exact test for the safety population for the Blinded Treatment Period (Period 2).

- Table RHBQ.6.7 defines the high and low baseline values as well as the limits that are specified as treatment-emergent and follow-up emergent. Note that weight does not have an abnormal baseline; therefore, the treatment-emergent and follow-up emergent values are determined by change from baseline.
- All post-baseline scheduled, unscheduled and repeated measurements at the same visit will be included.
- To assess increases, change from the maximum value during the baseline period or baseline to the maximum value during each study period will be used.

- To assess decreases, change from the minimum value during the baseline period or baseline to the minimum value during each study period will be used.
- For treatment-emergent high and low:
    - A treatment-emergent high result is defined as a change from a value less than or equal to the high limit at baseline to a value greater than the high limit at any time that meets the specified change criteria during the respective treatment periods.
    - A treatment-emergent low result is defined as a change from a value greater than or equal to the low limit at baseline to a value less than the low limit at any time that meets the specified change criteria during the respective treatment periods.
- For follow-up emergent high and low:
    - A follow-up emergent high result is defined as a change from a value less than or equal to the high limit at baseline to a value greater than the high limit at any time that meets the specified change criteria during the Post-Treatment Follow-up period.
    - A follow-up emergent low result is defined as a change from a value greater than or equal to the low limit at baseline to a value less than the low limit at any time that meets the specified change criteria during the Post-Treatment Follow-up period.

Table RHBQ.6.7. Categorical Criteria for Abnormal Treatment-Emergent Blood Pressures and Pulse Measurement, and Categorical Criteria for Weight Changes for Adults

Parameter	Low	High
Systolic BP (mm Hg) [a] (supine or sitting – forearm at heart level)	≤90 and decrease from baseline ≥20	≥140 and increase from baseline ≥20
Diastolic BP (mm Hg) [a] (supine or sitting – forearm at heart level)	≤50 and decrease from baseline ≥10	≥90 and increase from baseline ≥10
Pulse (bpm) [a] (supine or sitting)	<50 and decrease from baseline ≥15	>100 and increase from baseline ≥15
Weight (kg)	(Loss) decrease from baseline ≥7%	(Gain) increase from baseline ≥7%

Abbreviations: BP = blood pressure; bpm = beats per minute; kg = kilogram; mm Hg = millimeters of mercury.

[a] Baseline abnormal values are defined by the value presented.

6.15.6. Quick Inventory of Depressive Symptomatology–Self Report 16items (QIDS-SR16)

The QIDS-SR16 is a self-administered 16-item instrument intended to assess the existence and severity of symptoms of depression as listed in the American Psychiatric Association's (APA's) *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition* (DSM-IV) (APA 1994). The QIDS-SR16 scale is used to assess the potential impact of treatment on new onset or changes in depression, thoughts of death, and/or suicidal ideation severity. A patient is asked to consider each statement as it relates to the way they have felt for the past 7 days. There is a 4-point scale for each item ranging from 0 to 3. The 16 items corresponding to 9 depression domains are summed to give a single score ranging from 0 to 27, with higher scores denoting greater symptom severity. Additional information and the QIDS-SR16 questions may be found at the University of Pittsburgh IDS/QIDS resource page (2015 [WWW]).

The 9 domains assessed by the instrument are defined as:

1) Sleep disturbance (initial, middle, and late insomnia or hypersomnia): the highest score recorded for the four sleep items: #1 (falling asleep), #2 (sleep during the night), #3 (waking up too early) and #4 (sleeping too much). This domain is missing if all items are missing.
2) Sad mood: Item #5 (feeling sad). This domain is missing if the item is missing.
3) Decrease/increase in appetite/weight: the highest score recorded for the appetite/weight items: #6 (decreased appetite), #7 (increased appetite), #8 (decreased weight within the last two weeks), and #9 (increased weight within the last two weeks). This domain is missing if all items are missing or not applicable.
4) Concentration: Item #10 (concentration / decision making). This domain is missing if the item is missing.
5) Self-criticism: Item #11 (view of myself). This domain is missing if the item is missing.
6) Suicidal ideation: Item #12 (thoughts of death or suicide). This domain is missing if the item is missing.
7) Interest: Item #13 (general interest). This domain is missing if the item is missing.
8) Energy/fatigue: Item #14 (energy level). This domain is missing if the item is missing.
9) Psychomotor agitation/retardation: The highest score recorded for the two psychomotor items: #15 (feeling slowed down) and #16 (feeling restless). This domain is missing if all items are missing.

The QIDS-SR16 total score is the sum of the above domain scores. The total score will be missing if any domain score is missing.

The QIDS-SR16 total scores are categorized as follows:

- None (no depression): 0 – 5

- Mild: 6 – 10
- Moderate: 11 – 15
- Severe: 16 – 20
- Very severe: 21 – 27.

The following summaries will be produced for QIDS-SR16 total score category by treatment group for safety population during the Blinded Treatment Period (Period 2):

- The number and percentage of patients falling into each QIDS-SR16 total score category at each scheduled visit.
- Shift from maximum baseline to each post-baseline visit in QIDS-SR16 total score category.
- The number and percentage of patients falling into the following categories based upon the maximum post-baseline QIDS-SR16 total score:
    - Improved: maximum post-baseline category < maximum baseline category.
    - Worsened: maximum post-baseline category > maximum baseline category.
    - Same: maximum post-baseline category = maximum baseline category.

In addition, the number and percentage of patients falling into the following groups based upon the maximum post-baseline QIDS-SR16 item 12 (Thoughts of Death or Suicide) score will be summarized by treatment group for safety population during the Blinded Treatment Period (Period 2):

- Improved: maximum post-baseline QIDS-SR16 item 12 score < maximum baseline item 12 score.
- Worsened: maximum post-baseline QIDS-SR16 item 12 score > maximum baseline item 12 score.
- Same: maximum post-baseline QIDS-SR16 item 12 score = maximum baseline item 12 score.

6.15.7. Columbia-Suicide Severity Rating Scale (C-SSRS)

The Columbia-Suicide Severity Rating Scale (C-SSRS) is an assessment tool that evaluates suicidal ideation and behavior. Information on the C-SSRS scale can be found through the following link: http://www.cssrs.columbia.edu.

Specifically, the following outcomes are C-SSRS categories and have binary responses (yes/no). The categories have been re-ordered from the actual scale to facilitate the definitions of the composite and comparative endpoints, and to enable clarity in the presentation of the results.

Category 1 – Wish to be Dead
Category 2 – Non-specific Active Suicidal Thoughts

Category 3 – Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4 – Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5 – Active Suicidal Ideation with Specific Plan and Intent
Category 6 – Preparatory Acts or Behavior
Category 7 – Aborted Attempt
Category 8 – Interrupted Attempt
Category 9 – Actual Attempt (non-fatal)
Category 10 – Completed Suicide

Self-injurious behavior without suicidal intent is also a C-SSRS outcome (although not suicide-related) and has a binary response (yes/no).

Composite endpoints based on the above categories are defined below.

- Suicidal ideation: A "yes" answer at any time during treatment to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS.
- Suicidal behavior: A "yes" answer at any time during treatment to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS.
- Suicidal ideation or behavior: A "yes" answer at any time during treatment to any one of the ten suicidal ideation and behavior questions (Categories 1-10) on the C-SSRS.

Given that few or no suicidal ideation or behaviors are anticipated, 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 (i.e., if a patient 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.

Note that missing data should not be imputed.

6.16. Immunogenicity

6.16.1. Definitions and Terms

The following sample- and patient-related definitions and parameters will be used to describe the immunogenicity data.

6.16.1.1. Sample Category Definitions

Samples are classified into the following categories:

- **Unevaluable sample:** Sample could not be tested for ADA due to sample loss, mishandling, or errors in collection, processing, storage, etc.
- **Anti-drug antibody (ADA) Positive sample:** The presences of ADA is detected and confirmed. The samples are reported as positive. If the sample is positive, a titer value is reported.

- **Neutralizing anti-drug antibody (NAb) Positive sample:** NAb are reported as detected.
- **Antidrug antibody (ADA) Negative sample:** The presence of ADA is not detected and the assay drug tolerance level is not exceeded.
- **NAb Negative sample:** The presence of NAb is not detected and the assay drug tolerance level is not exceeded.
- **Inconclusive sample:** when ADA/NAb is not detected in a sample but drug is present in the same sample at a level that can cause interference in the ADA/NAb detection method, then the negative ADA/NAb result cannot be confirmed and the sample should be considered inconclusive.
    - Confirmation of a negative ADA result is based on observed/measured ixekizumab levels. Ixekizumab concentration allows for confirmation of negative NAb results.

Figure RHBQ.6.1 illustrates the relationship of some of the above terms.

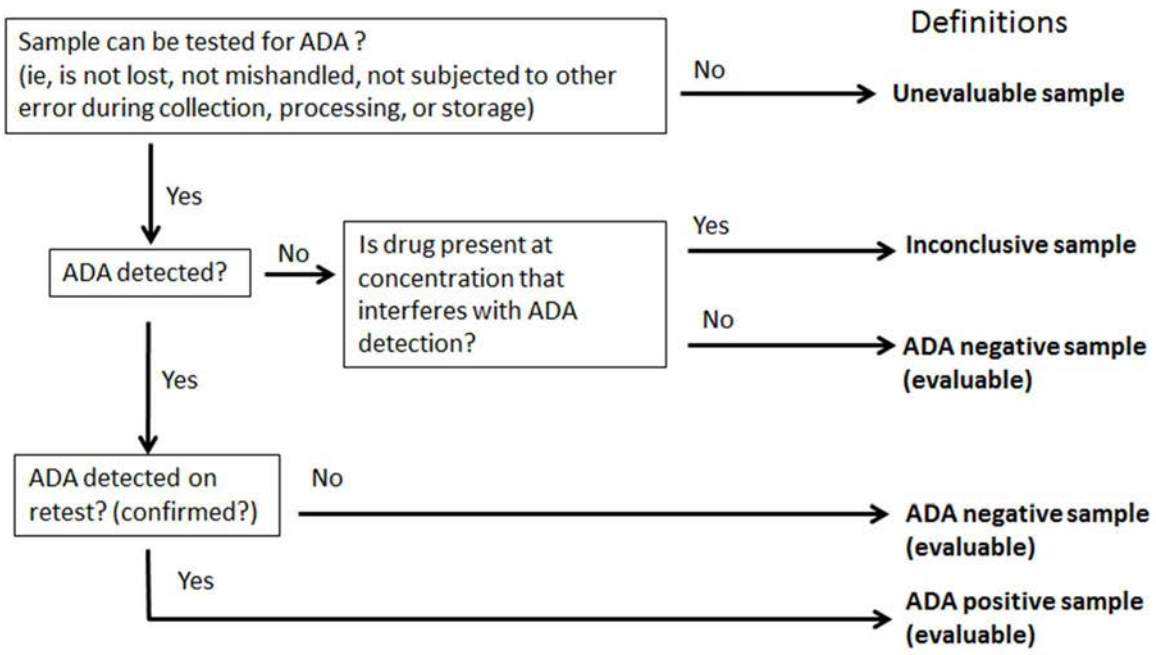


Figure RHBQ.6.1. Sample definitions.

6.16.1.2. Patient Category Definitions

The following categories are applied to patients based on the classification of their samples:

- **Unevaluable patient:**

a) a patient with no evaluable baseline sample and/or no evaluable post-baseline samples or

b) a patient with an evaluable baseline sample but no evaluable post-baseline sample or

c) a patient with no evaluable baseline sample, but whose evaluable post-baseline values are all ADA positive or a mix of positive and negative. (Note: If all post-baseline samples are negative, the patient is considered 'evaluable' and will be classified as ADA-negative.)

- **Evaluable patient:**

a) a patient with an evaluable baseline sample and at least 1 evaluable post-baseline sample (that is, sample after administration of study drug) or

b) a patient with no evaluable baseline sample whose evaluable post-baseline samples are all ADA negative.

Figure RHBQ.6.2 illustrates the relationship of the above terms.

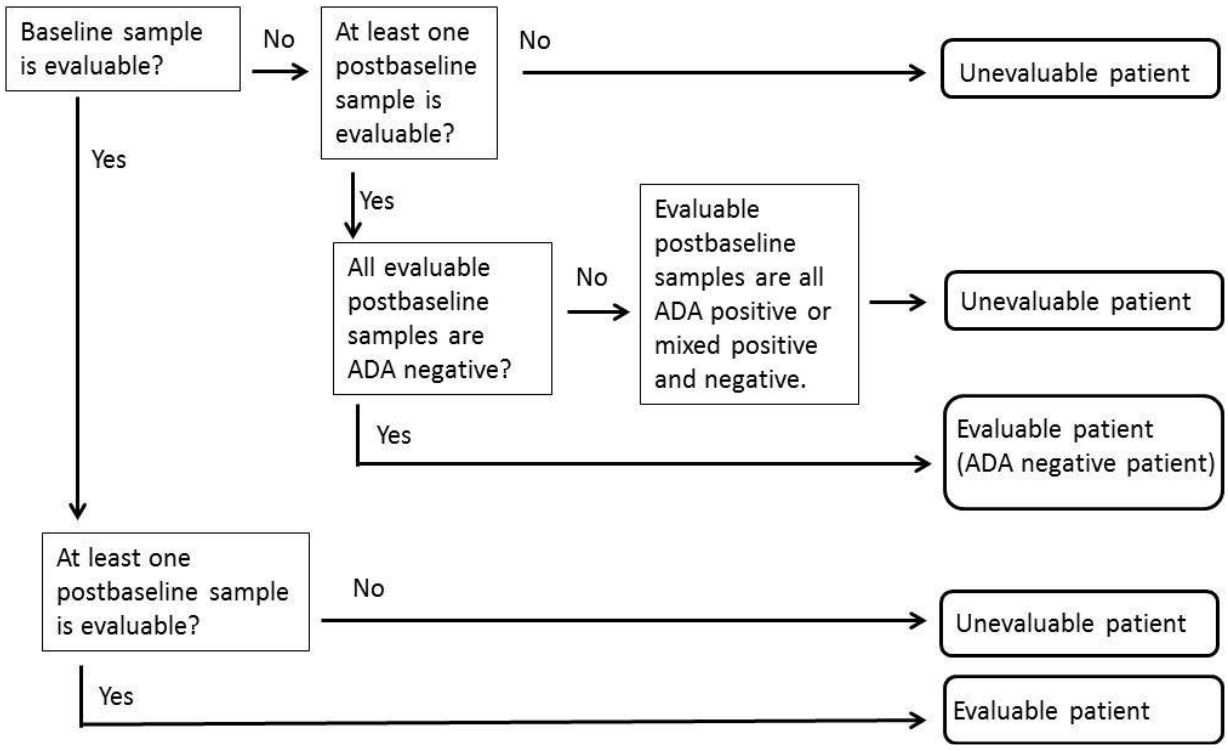


Figure RHBQ.6.2. Patient categories (evaluable/unevaluable) based on sample status at baseline and post-baseline.

6.16.1.3. Definitions for Clinical Interpretation of Assay Results

- **Baseline:** For immunogenicity analyses during Period 2, baseline is the last non-missing observation on, or prior to, the date of the first injection of study treatment of ixekizumab

(Week 0). Unless otherwise specified, the baseline for subsequent treatment periods is defined as the last non-missing observation on, or prior to, the date of first injection of ixekizumab. For patients originally randomized to ixekizumab during Period 2, baseline is the last non-missing observation on, or prior to, the date of the first injection of study treatment for Period 2 (Week 0). For patients who are not originally randomized to ixekizumab in Period 2, baseline is the last non-missing observation on, or prior to, the date of the first injection of ixekizumab. See Table RHBQ.6.8 for further details.

Table RHBQ.6.8. Baseline Definition for Immunogenicity Analyses for Extended Treatment Period

Treatment Assignment for Blinded Treatment Period (Period 2)	Treatment Assignment for Open-Label Treatment Period (Period 3)	Baseline for Open-Label Treatment Period Analysis[a]
Ixekizumab	Ixekizumab	Week 0
Placebo	Ixekizumab	Week 12

[a] Last non-missing observation on, or prior to, the date of the first injection of study treatment at the defined week.

- **Baseline ADA positive (pre-existing antibody):** ADA detected in a sample collected at baseline.
- **TE-ADA positive:** a) a patient with a ≥4-fold increase over a positive baseline antibody titer (Tier 3); or b) for a negative baseline titer, a patient with an increase from the baseline to a level of ≥1:10.
- **Baseline ADA-negative:** ADA is not detected in a sample collected at baseline.
- **TE-ADA inconclusive patient:** A patient without a TE-ADA positive sample and with at least one sample for which drug levels may interfere with the ADA assay.
- **TE-ADA negative patient:** A patient who is evaluable for TE-ADA and is not either TE-ADA positive or TE-ADA inconclusive.

Figure RHBQ.6.3 illustrates the relationship of some of these terms.

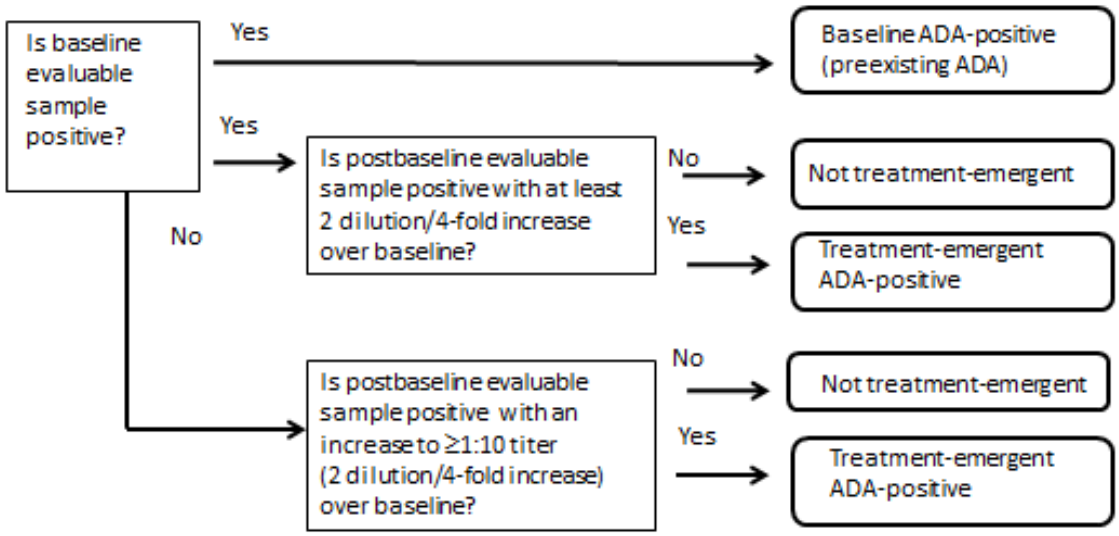


Figure RHBQ.6.3. Relationship of terms for clinical interpretation of assay results for evaluable patients.

- **Incidence of TE-ADA:** Patients with TE-ADA as a proportion of the evaluable patient population during the treatment period. This excludes unevaluable patients.
- **Follow-up emergent ADA:** ADA is first detected during the follow-up period, after study drug administration is discontinued. This category includes patients negative at baseline who increased to ≥1:10 titer (4-fold increase/2 dilutions) after baseline in the follow-up period or patients ADA positive at baseline and increased at least 4-fold (2 dilutions) over baseline for the first time in the follow-up period.
- **Incidence of follow-up emergent ADA:** Patients with follow-up emergent ADA as a proportion of the follow-up evaluable patient population. This excludes unevaluable patients.

All ADA positive samples will be evaluated for NAb. Definitions for NAb patient status will be defined as follows:

- **NAb-positive patient:** A patient where a NAb positive result is detected for ≥1 TE-ADA positive samples.
- **NAb-inconclusive patient:** A patient without a NAb positive sample and with at least one sample for which drug levels may interfere with the NAb assay.
- **NAb-negative patient:** A patient who is evaluable for NAb and is not either NAb positive or NAb inconclusive.

A flow chart that reflects the connection between the analytical test results and the clinical interpretation based on the definitions is shown in Figure RHBQ.6.4.

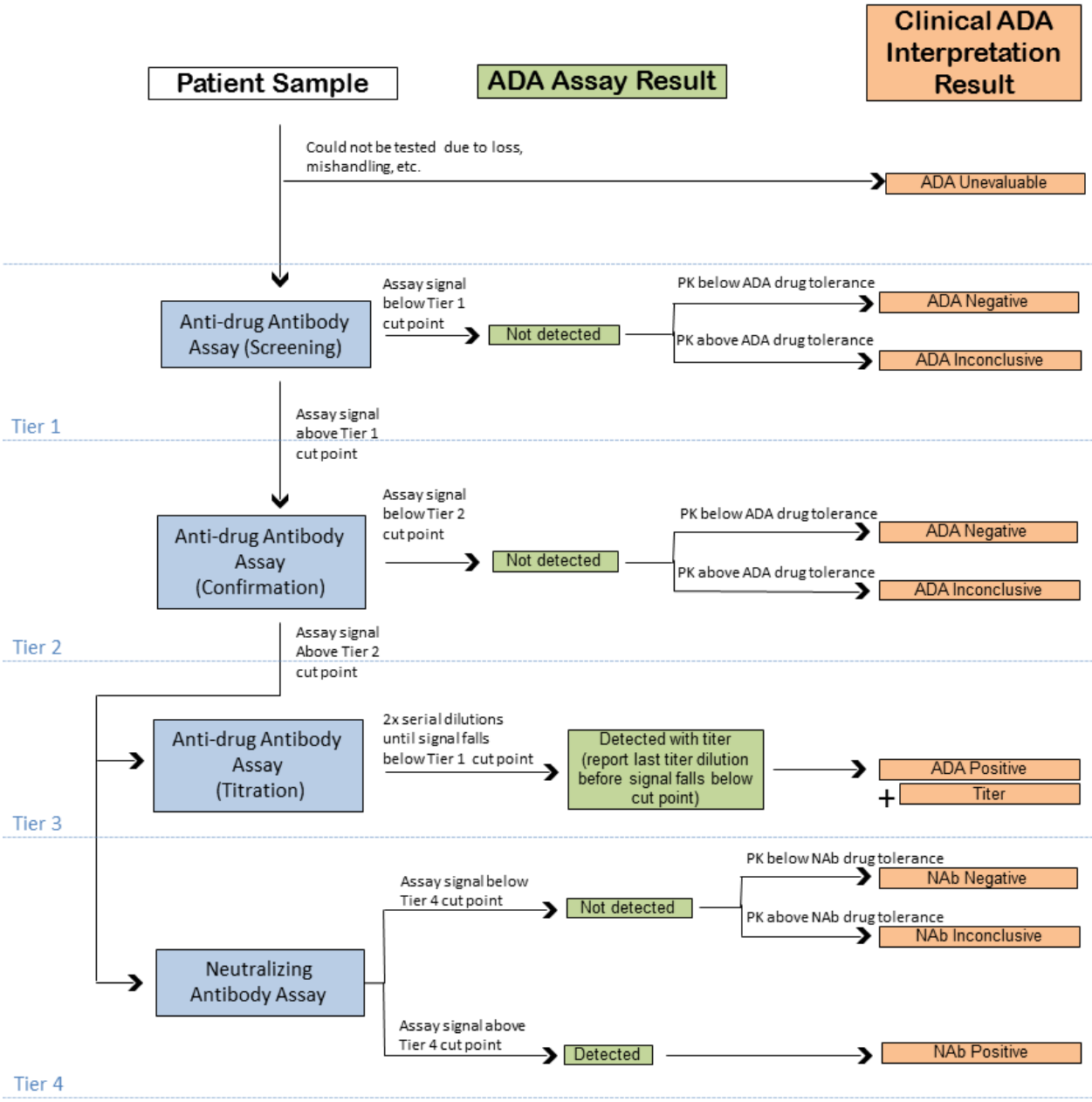


Figure RHBQ.6.4. Flow chart of ADA assessment with clinical interpretation of the various result possibilities.

6.16.2. Immunogenicity Analyses

Immunogenicity evaluable patients will be identified as TE-ADA positive, TE-ADA negative, or TE-ADA inconclusive, according to the definitions provided in Section 6.16.1.2 and further grouped into TE-ADA status groups and time-varying TE-ADA status groups:

TE-ADA Status Groups:

- TE-ADA status (positive, negative, or inconclusive);

- NAb status (positive, negative, or inconclusive) for TE-ADA positive patients; and
- TE-ADA titer groups for TE-ADA positive patients:
    - Low Titer: TE-ADA titer value (last observation carried forward [LOCF]) < 1:160;
    - Moderate Titer: TE-ADA titer value (LOCF) ≥1:160 and < 1:1,280; and
    - High Titer: TE-ADA titer value (LOCF) ≥1:1,280.

The LOCF approach will only be performed in the section of immunogenicity analyses. This approach is identical to the mBOCF approach, with one exception: for patients discontinuing investigational product due to an AE, the last non-missing post-baseline observation before discontinuation will be carried forward to the corresponding endpoint for evaluation. Randomized patients without at least one post-baseline observation will not be included for evaluation.

Time-varying TE-ADA Status Groups:

Individual ADA samples will be ascribed into three different dichotomous variables as explained in Table RHBQ.6.9. Each variable has possible values of a 'greater-TE-ADA status' or a 'lesser-TE-ADA status', in the sense that the level of TE-ADA detected in the greater-TE-ADA category is higher than in the lesser-TE-ADA category.

Table RHBQ.6.9. TE-ADA Status Dichotomous Variables for AE Analysis

TE-ADA Status Dichotomous Variable	Greater-TE-ADA Status	Lesser-TE-ADA Status
TE-ADA titer	TE-ADA titer	not TE-ADA titer
TE-ADA moderate-to-high	TE-ADA positive with moderate titer or high titer	not TE-ADA positive, or TE-ADA positive with low titer
TE-ADA high status	TE-ADA positive with high titer	not TE-ADA positive, or TE-ADA positive with low or moderate titer

Abbreviations: AE = adverse event; TE-ADA = treatment-emergent anti-drug antibody.
Note: for purpose of this analysis, TE-ADA Inconclusive is taken to be "not TE-ADA positive".
Note: a TE-ADA low is defined as a TE-ADA positive with a titer value < 1:160; a TE-ADA moderate is defined as a TE-ADA positive with a titer value ≥1:160 and < 1:1,280; and a TE-ADA high is defined as a TE-ADA positive with a titer value ≥1:1,280.

For each TE-ADA status dichotomous variable, a time-varying TE-ADA status will be computed. At time *t* the TE-ADA status is taken to be the highest of the TE-ADA values bracketing time *t*. More formally, the TE-ADA status at time *t* is given by the greater of (a) the TE-ADA status at the most-recent post-baseline measurement prior to *t*, and (b) the TE-ADA status at the first TE-ADA post-baseline measurement at or after time *t*. In this computation, 'greater' is given by the greater-TE-ADA status of Table RHBQ.6.9. If there is no value

satisfying criterion (a), then the value (b) is used. Similarly, if there is no value (b), then the value (a) is used.

For each TE-ADA status dichotomous variable, patients will be categorized according to whether they were (i) always in lesser-TE-ADA status post-baseline or (ii) at some point post-baseline, were in greater-TE-ADA status.

6.16.2.1. Analyses of Characteristics of ADA Immune Response

The analyses of ADA effects will be conducted on all evaluable patients within the defined safety population for the Blinded Treatment Period (Period 2), Open-label Treatment Period (Period 3) and within the defined follow-up population for the Post-Treatment Follow-up Period (Period 4).

The overall frequency and percentage (incidence) of patients will be summarized for the TE-ADA status groups and the time-varying TE-ADA status groups. Scheduled visits, unscheduled visits, and repeat measurements will be included.

The time to the development of TE-ADAs (TE-ADA positive, low titer, moderate titer, high titer, and NAb positive) will be calculated as follows:

> *Time to development of TE-ADAs/NAb (in weeks) = (Date of development of TE-ADAs/NAb – Date of first injection of study treatment + 1) / 7.*

If a patient has not developed TE-ADAs/NAbs, they will be censored at the date of the last immunogenicity assessment. If a patient does not have any post-baseline assessments for immunogenicity, they will be censored at the date of randomization.

Descriptive statistics, including 25th percentile, 50th percentile (median), 75th percentile, and corresponding 95% confidence intervals as well as probability of TE-ADA/NAb positive by endpoint summarized by treatment group, will also be provided if sufficient data is present. A Kaplan-Meier plot of the time to development of treatment-emergent ADA/NAb will be presented by treatment group, also if sufficient data is present. The log-rank test will be used to test the null hypothesis against the alternate hypothesis that the time to TE-ADA/NAb is not equal between each ixekizumab dose and placebo. Caution should be exercised in the interpretation of time-to event analyses, and related statistics, given the limited sampling scheme for immunogenicity testing.

For each TE-ADA status dichotomous variable (as defined in Table RHBQ.6.9), summaries will be provided of the total post-baseline time in the greater-TE-ADA status for patients who were at some point post-baseline in the greater-TE-ADA status group. Post-baseline time in greater-TE-ADA status for each patient will be aggregated.

A by-patient listing to include treatment, visit date, visit, ADA result, TE-ADA result, NAb result, ADA titer value, ixekizumab concentration, ADA and NAb inconclusive results will also be provided for the individual studies, for patients with any one sample of ADA (or NAb) positive or inconclusive.

6.16.2.2. Analyses of Treatment-emergent ADA Effects on Efficacy

Efficacy analyses for the Blinded Treatment Period (Period 2) and Open-label Treatment Period (Period 3) will be conducted on all evaluable patients within the ITT Population.

Analyses will be performed to examine how patient TE-ADA effects sPGA of Genitalia (0,1) at week 12 with NRI by the TE-ADA status groups as described in Section 6.16.1.2. Note that the TE-ADA negative category will also be included in the NAb status and TE-ADA titer group analyses.

For sPGA of Genitalia (0,1) at Week 12, a logistic regression model with treatment group, TE-ADA status and the interaction of treatment group-by-TE-ADA status included as factors will be used to test the interaction of treatment group-by-TE-ADA status. The p-value associated with the interaction term will be used to assess if the treatment groups effect is consistent across TE-ADA status groups. When the interaction term is statistically significant, the association between responder status and treatment groups depends, in some manner, on the status. The interaction will be tested at the 10% significance level. Treatment group differences will be evaluated within each subgroup using Fisher's exact test regardless of whether the interaction is statistically significant.

6.16.2.3. Analyses of Treatment-emergent ADA on Specific Adverse Events

The analyses of ADA effects on safety will be conducted on all evaluable patients within the defined safety population for Period 2 and Open-Label Treatment Population for Period 3.

Adverse events of special interest of allergic reaction/hypersensitivity (anaphylaxis and non-anaphylaxis) and of injection-site reactions will be included in an assessment of AE to TE-ADA over time. See Section 6.15.3.1 for the definitions of the AESIs. Timing of an AE will be taken to be the reported AE start date.

For each TE-ADA status dichotomous variable (as defined in Table RHBQ.6.9), patients will be categorized according to whether they were (i) always in lesser-TE-ADA status post-baseline or (ii) at some point post-baseline, were in greater-TE-ADA status. For each AESI, within the time-varying TE-ADA status groups, a summary will be provided of the number of patients who had no event, events only while in lesser-TE-ADA status for group (i), or – for group (ii) – at least one event while in greater-TE-ADA status.

Additionally, summaries will be provided of the total number of AESI events (with unique start dates) by time-varying TE-ADA status groups at the event date. The summaries will aggregate time respectively in greater-TE-ADA status and in lesser-TE-ADA status, along with the event rates (rates per 100 patient-years) relative to those aggregate times.

By-patient listings will be provided of patients with TE-ADA who experience a treatment-emergent allergic reaction/hypersensitivity reaction or an injection site reaction.

6.17. Subgroup Analyses

6.17.1. Efficacy Subgroup Analyses

Subgroup analysis will be conducted for the primary endpoint at Week 12 using the ITT Population for the Blinded Treatment Period (Period 2).

- **The proportion of patients achieving an sPGA of Genitalia (0,1) (primary endpoint):** a logistic regression model with treatment, subgroup, and the interaction of treatment-by-subgroup as factors will be used. The treatment-by subgroup interaction will be tested at the significance level of 0.10. Treatment group differences will be evaluated within each category of the subgroup using Fisher's exact test, regardless of whether the interaction is statistically significant. Missing data will be imputed using NRI, as described in Section 6.3.1. If any group within the subgroup is less than 10% of the total population, only descriptive statistics will be provided for that subgroup (that is, no inferential testing will be done).

The following subgroups will be analyzed:

- BSA category (<10%, ≥10%)
- Sex (male, female)
- Age category (<40 years, ≥40 years)
- Geographic Region (US [including Puerto Rico], non-US)

6.17.2. Safety Subgroup Analyses

Safety subgroup analysis for common TEAEs and AESI of allergy reaction/hypersensitivity and infections will be summarized by treatment group and overall, using the Safety Population for the Blinded Treatment Period (Period 2). The common TEAEs will be presented by MedDRA PT within SOC. The AESI of allergy reaction/hypersensitivity will be presented by anaphylaxis and non-anaphylaxis events, by PT within category. The AESI of infection will be presented by PT.

A logistic regression model with treatment group, subgroup, and the interaction of subgroup-by-treatment group as factors will be used. The subgroup-by-treatment group interaction will be tested at the significance level of 0.10. The response variable will be each AE. Treatment group differences will be evaluated within each category of the subgroup using Fisher's exact test, regardless of whether the interaction is statistically significant. If any group within the subgroup is less than 10% of the total population, only the descriptive statistics will be provided for that subgroup (that is, no inferential testing will be done).

The following subgroups will be analyzed:

- BSA category (<10%, ≥10%)
- Sex (male, female)

- Age category (<40 years, ≥40 years)
- Geographic Region (US, non-US)

6.18. Interim Analyses

An interim (primary) database lock and the unblinding (of Lilly personnel) will occur and the interim analysis will be performed at the time (that is, a cut-off date) the last patient completes Visit 7 (Week 12) or ETV. This interim database lock will include all data collected up to the cut-off date including follow-up data from patients that have begun Post-Treatment Follow-up Period (Period 4). Because the study will still be ongoing for the Post-Treatment Follow-up Period at the time of this database lock, the analysis will be referred to as an interim analysis. This interim analysis will include the final analysis for the Blinded Treatment Period (Period 2) of the study; therefore, there will be no alpha adjustment due to this interim analysis.

A final database lock will occur after the Post-Treatment Follow-up Period is completed.

6.19. 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 summary of AEs, provided as a dataset which will be converted to an XML file. Both SAEs and 'Other' AEs are summarized: by treatment group, by MedDRA preferred term.

- An AE is considered 'Serious' whether or not it is a treatment emergent adverse event (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/subjects 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 such as the CSR.

6.20. Exploratory Analyses

Exploratory analyses other than those mentioned in Table RHBQ.6.5 and Section 6.20.1 may be added to future versions of this analysis plan, as deemed appropriate.

6.20.1. Psychometric Analyses of Diary Endpoints and sPGA of Genitalia

6.20.1.1. Psychometric Evaluation Overview

Psychometric evaluation is an iterative process beginning with qualitative assessment and also involving quantitative evaluation, wherein item and scale quality are examined to eliminate poorly performing or redundant items, to assess the scale structure and content validity, and to evaluate the measure's relationship to previously utilized and validated measures with regard to the constructs that the instrument is intended to measure. This section describes analyses that will be conducted to evaluate the psychometric properties of the GPSS, GPSIS, SFQ, and sPGA of Genitalia. No single test can be used to determine the psychometric qualities of an instrument; instead, a series of tests are proposed, each designed to evaluate a different aspect of the instrument's performance. Upon completion of the psychometric analyses, an overall summary regarding the psychometric properties of the instrument will be made, based on the analysis results. The following analyses will be conducted for each instrument where possible:

- Sample and item descriptive analyses, item level evaluation, and scaling, which will include descriptive item statistics, floor and ceiling effects.
- For measures with multiple items (for example, GPSS) the dimensionality of the measure will be explored using factor analyses.
- Assessment of the measurement properties of the instruments, including reliability, validity and responsiveness.
- Instrument descriptive analyses for the daily diary, including evaluation of the reliability and validity of the recall periods.

6.20.1.2. Statistical Analysis

6.20.1.2.1. General Considerations

- All statistical tests will use a two-sided significance level of 0.05, unless otherwise noted. Statistical tests involving multiple comparisons (e.g., analysis of covariance [ANCOVA] models with multiple groups) will include Scheffe post-hoc tests which adjust for multiple comparisons and reduce the possibility of Type I errors.
- Mplus software Version 7.3 (Muthén and Muthén 1998–2012) will be used to conduct exploratory and confirmatory factor analysis.
- SAS statistical software will be used for the remaining psychometric analyses. All SAS codes that are prepared for these analyses will be validated in line with the expectations of the Food and Drug Administration (FDA) and European Medicines Agency (EMA).
- The scoring of all measures used in the analyses (for example, DLQI, SF-36) will be provided by the developers of the instruments and documented elsewhere in the SAP.

- Most analyses outlined below will be conducted using the ITT population. Unless stated otherwise, the analyses involving the subscale of GPSIS, Sexual Activity Avoidance Subscale, will be restricted to patients who answered "Yes" or "No due to my genital psoriasis" in GPSIS Item #1.

6.20.1.2.2. Descriptive Statistics

Descriptive statistics (sample size, mean, SD, minimum, median, maximum, and % missing) will be calculated for the diary endpoints, including total score and item score and relevant subscale, and sPGA of Genitalia at baseline and Week 12.

The frequency distributions for item responses will be calculated for each item within the GPSS, each subscale within the GPSIS, and each item within the SFQ, as well as for the sPGA of Genitalia at baseline and Week 12.

6.20.1.2.3. Factor Analysis

Exploratory Factor Analysis (EFA) will be conducted using baseline GPSS data. Exploratory Factor Analysis will be used to uncover the underlying factor structure of the measures. Screen plots and corresponding eigenvalues will be examined to empirically inform the number of factors underlying the GPSS items. Factor loadings, standardized mean square residual (SRMR), and root mean square error of approximation (RMSEA) will be examined to evaluate model goodness-of-fit. It is recommended that the SRMR and RMSEA values be less than 0.8. Orthogonal and/or oblique rotation (Yates 1987) will be used in the analysis since the factors are expected to be correlated. An approximate simple structure will be the criterion for accepting a factor solution; oblique rotation will be conducted and correlations between factors will be reported, if necessary.

The GPSS includes multiple items and these analyses will evaluate the performance of each item in order to identify items that may be redundant or perform poorly. Items with low standardized loading <0.40, or items with similar factor loadings across multiple domains indicate items that may be candidates for removal. Interpretability of factors and clinical judgment will also be considered in the selection of the final model.

Confirmatory factor analysis (CFA) will be conducted using Week 1 data to evaluate the conceptual framework for the GPSS. It is hypothesized that the GPSS will yield a one factor solution similar to other symptom measures for plaque psoriasis (Revicki et al. 2014). Confirmatory factor analysis will be conducted to model factor loadings based on the final conceptual model derived in the EFA analyses.

6.20.1.2.4. Differential Item Functioning (DIF)

Differential Item Functioning (DIF) analysis refers to procedures that assess whether items on an instrument perform in the same manner for various groups of patients. Differential Item Functioning procedures typically control for overall between-group differences on a criterion, usually the scores of the instrument (Zumbo 1999). In this study, DIF analysis will be conducted on the GPSS items score, GPSIS subscale and SFQ item to clarify whether the items perform differently by sex (male vs. female), while controlling for their scale scores. The scale scores as

determined on the basis of the factor analysis and Rasch model analysis will be used in the DIF analysis. Sample sizes of the subgroups will be checked prior to conducting DIF. Differential Item Functioning analysis will only be performed when sample sizes are sufficient (that is, ≥50).

Ordinal logistic regression models will be used in the DIF analyses. Differential Item Functioning will be conducted for each item. Each model will have the scale item score as the dependent variable, the grouping variable "sex" as the factor, the Rasch logit scale score as the covariate, and the interaction term for "sex" by Rasch logit scale score. The grouping variable "sex" is called a bias factor. The Rasch logit scale score is a derived score of the Rasch based on the factor analysis model and will be used as the measure of the underlying construct of severity for the DIF analysis. Ideally, only the underlying construct influences the way participants respond to the items, and the bias factor does not influence the dependent variable. However, if an item exhibits DIF, then the bias factor, in addition to the latent construct, also affects the way the participants respond to the items.

6.20.1.2.5. Psychometric Analyses

Following the examination of the distributional characteristics of measures, the following psychometric properties of the instruments will be evaluated: internal consistency (GPSS only), test-retest reliability, convergent validity, known-groups validity and sensitivity to change (responsiveness) using anchor- and distribution-based methods.

6.20.1.2.5.1. Internal Consistency

Internal consistency reliability addresses the extent to which individual items in the instrument measure a common underlying concept by examining the item-item or item-total, correlations for the GPSS using Cronbach's coefficient alpha (Cronbach 1951).

The values are presented descriptively on an interval level scale from 0 to 1.0, with higher scores indicating a more reliable (homogeneous) instrument. Cronbach's alpha coefficients between 0.7 and 0.9 will indicate good internal consistency, values between 0.4 and <0.7 will indicate moderate internal consistency, and values <0.4 will indicate low internal consistency reliability (Nunnally and Bernstein 1994, Cronbach 1951). Cronbach's alphas greater than 0.70 are generally considered acceptable for group comparisons (Hays and Revicki 2005). Internal consistency reliability will be assessed for the GPSS at the day prior to the first study drug injection and again at the day prior to the Week 12 injection based on item-level data. Cronbach's alpha for the GPSS total score and a summary of Cronbach's alpha levels with each constituent item deleted will also be reported.

Pearson and Spearman correlations will be used to calculate inter-item correlations for GPSS items. These correlations examine the extent to which items within the measure related to each other. These correlations will be used to confirm that the measures are performing in this sample as expected. Items will be flagged as being potentially redundant/overlapping when the item-item correlation is greater than 0.80.

6.20.1.2.5.2. Test-Retest Reliability

Test-retest reliability has been emphasized by the FDA as an important aspect of reliability in the FDA PRO Guidance (FDA 2009). Test-retest reliability reflects the ability of the instrument to give reproducible results when the clinical state is stable and indicates the degree to which a measure produces consistent results over several administrations. Test-retest reliability will be examined for: 1) GPSS items and total score, 2) GPSIS subscales, 3) SFQ item 2, and 4) sPGA of Genitalia prior to the first study drug injection. Stable subjects will be defined in two ways: 1) patients who had the same rating on the overall sPGA at screening and baseline visit, and 2) patients whose change in mGPASI score at screening and baseline visit <0.5 SD of mGPASI at baseline. For daily diary, patients are required to enter start entering GPSS at least 14 days prior to the first injection. The test score will be the average of 8 to 14 days prior to the first injection; whereas the re-test score will be the average of 1 to 7 days prior to the first injection, the same as baseline score. When there are less than 4 non-missing assessments within respective window, the search window will be expanded repeatedly until 4 non-missing assessments are available to compute the average score. Otherwise, the test, or re-test score will be set to missing. For weekly diaries, including GPSIS and SFQ, patients are also required to have 2 assessments collected prior to the first study drug injection in order to be included in the analysis.

Test-retest reliability will be assessed using intra-class correlations coefficients (ICCs) and paired sample t-tests among stable patients only.

Intra-class correlations coefficients range from 0 to 1.0, with higher scores indicating a more stable instrument. The hypothesis is that there will be no significant differences in scale scores when there is no change in disease status. Given that the majority of evaluated measures are single-item scales, a threshold of >0.60 will be supportive of test-retest based on the lower limit provided by Devellis (1991).

6.20.1.2.5.3. Construct Validity

6.20.1.2.5.3.1. Convergent and Divergent Validity

Construct validity is the degree to which a measure is related to other measures or constructed in a manner that is consistent with theory. *Convergent validity* involves demonstrating that different measures of the same concept substantially correlate, while *divergent* validity demonstrates that concepts that are supposed to be unrelated are, in fact, unrelated. The relationship between the 1) GPSS items and total score, 2) GPSIS subscales, 3) SFQ item 2, and 4) sPGA of Genitalia, and the DLQI domains, SF-36 scores, and mGPASI will be examined using Pearson, Spearman rank correlations, and polychoric correlation as appropriate.

Convergent validity of the scales of interest will be demonstrated at baseline and Week 12 by moderate to high correlations between the measures, except where noted. Moderate to high correlations (≥.30) demonstrating convergent validity (Cohen 1988) are expected for:

1. GPSS individual item and total scores and
    a. DLQI symptoms and feelings domain score, personal relationship score and total score

   b. SF-36 PCS scores
   c. mGPASI score
   d. PatGA-Genital
2. GPSIS subscales and
   a. DLQI personal relationships domain score
   b. DLQI sexual difficulties item
   c. PatGA-Genital
3. SFQ item 2 and
   a. DLQI personal relationships domain score
   b. DLQI sexual difficulties item
   c. PatGA-Genital
4. sPGA of Genitalia
   a. DLQI symptoms and feelings domain score
   b. mGPASI
   c. PatGA-Genital

Divergent validity of the scale of interest will be assessed at baseline, and Week 12. Pearson, Spearman's rank correlation and polychoric correlation will be calculated between 1) GPSS individual items and total score, 2) GPSIS subscales, 3) SFQ item 2, and 4) sPGA of Genitalia, and the SF-36 mental health, role physical and role emotional scores, and the DLQI daily activities, leisure, and work and school domain scores. Small correlations (demonstrating discriminant validity) are anticipated for each of these comparisons (correlations < 0.30; Cohen 1988).

6.20.1.2.5.3.2. Known-Groups Validity

Known-group validity is the extent to which scores from an instrument are different for groups of participants that differ on a relevant clinical or other indicator. Determining known-group validity involves evaluating an instrument in relation to clinical measures of disease status (Stewart et al. 1992; Hays and Revicki 2005). Known-groups validity will be assessed at baseline and Week 12 for: 1) GPSS individual items and total score, 2) GPSIS subscales, 3) SFQ item 2, and 4) sPGA of Genitalia by disease severity. Groups will be defined using several different criteria: 1) using DLQI score categories (score categories: 2-10, 11-20, 21-30); and 2) using PatGA-Genitalia scores (score categories: 0-1, 2-3, 4-5. Additionally, for exploratory purposes, known-groups validity of the SFQ item 2 will be assessed at baseline-and Week 12 using the DLQI sexual difficulties item to define subgroups of patients. Mean scores on the individual items/subscale/total score will be compared for each subgroup using ANCOVA at baseline, and Week 12 controlling for age, sex, and baseline BSA (<10% vs ≥10%).

When the sample size in a subgroup is less than 10% of the analysis population, the subgroup will be repeatedly collapsed with the adjunct subgroups until the sample size in the collapsed subgroup is greater than 10% of the analysis population.

6.20.1.2.5.4. Responsiveness

Responsiveness refers to the extent to which the instrument can detect changes in patients known to have changed in clinical status (Hays and Revicki 2005). The responsiveness of the 1) GPSS item and total score, 2) GPSS subscales, 3) SFQ Item #2, and 4) sPGA of Genitalia will be determined by comparing changes in item/total scores with the changes in selected clinical and health status measures.

The following variables and time points will be used for these analyses:

- PatGA-Genitalia: The change scores of the GPSS item and total score from baseline to Week 12 will be correlated with changes in the PatGA-Genitalia at these time points. The change scores of the individual items and GPSS total score from baseline to Week 12 will also be examined using PatGA-Genitalia as an anchor. The means of GPSS and sPGA of Genitalia scores will be compared for patients reporting ≥1 point PatGA-Genitalia Improvement, no change, or ≥1 worsening in PatGA-Genitalia using ANCOVA adjusted for age, sex, genital psoriasis duration and baseline scores.
- mGPASI: The change scores of the GPSS individual item and total score from baseline to Week 12 will be correlated with changes in the mGPASI at these time points. For exploratory purposes, the change scores of the GPSS individual item and total score from baseline to Week 12 will also be examined for mGPASI defined groups using ANCOVA. Groups will be defined as patients who experienced improvement on the mGPASI of ≥0.5 SD versus improvement <0.5 SD.

Additionally, for exploratory purposes, responsiveness of the subscale of GPSIS, Sexual Activity Avoidance Subscale, and sPGA of Genitalia will be assessed using the Week 12 score. Both scales at Week 12 will be anchored on PatGA-Genitalia and mGPASI defined previously.

6.20.1.2.5.5. Clinical Significance

To ensure proper interpretation of the data, it is useful to develop evidence for clinical significance of score changes. The lower end of a range of clinical significance can be termed the responder definition (RD), which is defined as "the individual patient PRO score change over a predetermined time period that should be interpreted as a treatment benefit" (FDA 2009). In the absence of a single, unequivocal indicator of change, a number of analytic methods will be used to assess the RD of the GPSS individual item and total score, GPSIS, and SFQ. The results of these analyses can be used to guide score interpretation. Both anchor-based methods and distribution-based methods will be used for determining RD. However, anchor-based methods are preferred by the FDA for interpretation of PRO scores (FDA 2009) and will be considered the primary analysis. Distribution-based methods will be considered supportive and secondary (Revicki et al. 2008).

6.20.1.2.5.5.1. Anchor-Based Method

PatGA-Genitalia will be used as an anchor to determine the RD of each measure, including GPSS item and total score, GPSIS subscale, and sPGA of Genitalia. The RD will be estimated as the mean change scores of patients who improved by one and two points on the PatGA-Genitalia from baseline to Week 12. Effect size (ES) statistics will then be calculated for each group of patients as the mean difference between the baseline and Week 12 visit scores divided by the standard deviation of the baseline domain scores of each measure (Cohen 1988). The following cutoff values will be used to interpret effect size: small ES = 0.2, moderate ES = 0.5, and large ES = 0.8.

Additionally, the RD of the change in GPSS item and total score and the subscale of GPSIS, Sexual Activity Avoidance Subscale, from baseline to Week 12 will be derived by anchoring them on sPGA of Genitalia (0,1) at Week 12. The correlation between change in each measure and change in sPGA of Genitalia from baseline to Week 12 will first be computed to assess the strength of the anchor. A logistic model will then be utilized by regressing the change in each measure on the anchor. Concordance index will be computed to assess the adequacy of the model fitting. Predictive statistics, including sensitivity, specificity, positive predictive value, negative predictive value and Youden index, will be presented at each possible change of value for each measure.

A RD will be derived for sPGA of Genitalia and the subscale of GPSIS, Sexual Activity Avoidance Subscale at Week 12 by anchoring on 1) PatGA-Genitalia (0,1) at Week 12 as an anchor. Correlation between each measure and anchor variable will be first assessed prior to constructing the logistic model.

A RD will be derived for SFQ Item #2 and GPSIS subscales comprising the change of score and static score at Week 12 using PatGA-Genital (0,1) and sPGA of Genitalia (0,1) at Week 12 as anchors.

6.20.1.2.5.5.2. Distribution-Based Method

For the distribution-based methods, supportive evidence for the RD will be assessed using one standard error of measurement (SEM) and 0.5 SD of the scales of interest at baseline.

Standard of error of measurement has been proposed as a useful distribution-based statistic for evaluating clinically meaningful change in health-related quality of life (HRQL) measures (Wyrwich et al. 1999a; Wyrwich et al. 1999b). The SEM describes the error associated with the measure and is estimated by the baseline standard deviation of the measure multiplied by the square root of one minus its reliability coefficient (ICC from the test-retest assessment). Previous research suggests that one SEM is roughly associated with a clinically important difference for PRO measures (Wyrwich et al. 1999a; Guyatt et al. 2002). Other research suggests changes (or differences) of 0.20 to 0.30 effect size may be indicative of a minimal important difference (Osoba and King 2005; Revicki et al. 2008).

An alternative distribution-based approach consists of calculating a 0.5 SD of the scales of interest at baseline. It has been suggested that one-half of a standard deviation of a measure

represents a clinically meaningful change, while a change corresponding to 0.2 of a standard deviation of measure is a small effect (Norman and Streiner 2003). The 0.5 SD estimate can be considered to provide an upper boundary for what would constitute a meaningful change, while 0.2 provides a lower boundary (Revicki et al. 2008). Given that a clinically important change is of interest, a 0.5 SD will be used, with 0.2 SD calculated to provide reference for range.

The number and percentage of patients meeting the various RDs derived from anchor- and distribution-based methods for each measure will also be calculated.

Using the results from both the anchor- and distribution-based approaches, triangulation across possible thresholds indicating a minimal but meaningful change will be considered, and a reasonable RD will be selected (Leidy and Wyrwich 2005).

6.20.1.2.5.5.3. Cumulative Distribution Functions

Cumulative distribution function (CDF) shows the proportion of the population scoring less than or equal to each possible change score. Cumulative distribution functions are useful in that they graphically characterize the treatment effect or differences between groups. Cumulative distribution functions will be constructed for changes from baseline to Week 12 in GPSS individual items and total scores, and sPGA of Genitalia scores stratified by treatment and by responder status defined by PatGA-Genitalia. For all CDF plots, the x-axis is the score change from baseline to final visit for each measure. The y-axis is the cumulative proportion of the patients in each of the corresponding categories that reach the score change on the x-axis.

Another CDF will be presented for the subscale of GPSIS subscales and sPGA of Genitalia score at Week 12 stratified by treatment groups.

7. Unblinding Plan

A separate document will describe the blinding and unblinding plan.

8. References

[APA] American Psychiatry Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington DC: American Psychiatric Association; 1994.

Basra MKA, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. *Br J Dermatol*. 2008;159(5):997-1035.

Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc; 1988.

Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951;16(3):297-334.

DeVellis RF. Scale Development: Theory and Applications. Newbury Park, CA, USA: Sage, 1991.

[FDA] Food Drug Administration. Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. *Fed Regist*. 2009;74(235):65132-65133.

Diggle PJ, Liang KY, Zeger SL. Analysis of Longitudinal Data. Oxford: Oxford University Press; 1994.

Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)—a simple practical measure for routine clinical use. *Clin Exp Dermatol*. 1994;19(3):210-216.

Fredriksson T, Pettersson U. Severe psoriasis-oral therapy with a new retinoid. *Dermatologica*. 1978;157(4):238-244.

Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. Paper presented at: Mayo Clinic Proceedings; 2002.

Hays RD, Revicki D. Reliability and validity (including responsiveness). In: Fayers P, Hays RD, eds. Assessing Quality of Life in Clinical Trials: Methods and Practice. 2nd ed. New York, NY: Oxford University Press; 2005:25-39.

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.

Khilji FA, Gonzalez M, Finlay AY. Clinical meaning of change in Dermatology Life Quality Index scores. *Br J Dermatol*. 2002;147(suppl 62):50.

Leidy NK, Wyrwich KW. Bridging the gap: using triangulation methodology to estimate minimal clinically important differences (MCIDs). *COPD*. 2005;2(1):157-165.

Leonardi CL, Kimball AB, Papp KA, Yeilding N, Guzzo C, Wang Y, Li S, Dooley LT, Gordon KB; PHOENIX 1 study investigators. Efficacy and safety of ustekinumab, a human interleukin- 12/23 monoclonal antibody, in patients with psoriasis: 76-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 1). *Lancet.* 2008;371(9625):1665-1674. [Erratum appears in Lancet. 2008;371(9627):1838].

Muthén L, Muthén B. Mplus User's Guide. 7 ed. Los Angeles, CA: Muthén & Muthén; 1998-2012.

Norman GR, Streiner DL. *PDQ statistics.* Vol 1: PMPH-USA; 2003.

Nunnally JC, Bernstein I. *Psychometric theory.* 3 ed. New York,: McGraw-Hill; 1994.

Osoba D, King M. Meaningful differences. *Assessing quality of life in clinical trials.* 2005;2:243-257.

Papp KA, Langley RG, Lebwohl M, Krueger GG, Szapary P, Yeilding N, Guzzo C, Hsu M-C, Wang Y, Li S, Dooley LT, Reich K. Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 52-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 2). *Lancet.* 2008;371(9625):1675-1684.

Revicki DA, Jin Y, Wilson HD, Chau D, Viswanathan HN. Reliability and validity of the psoriasis symptom inventory in patients with moderate-to-severe psoriasis. *J Dermatolog Treat.* 2014;25(1):8-14.

Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J. Clin. Epidemiol.* 2008;61(2):102-109.

Sampson HA, Munoz-Furlong A, Campbell RL, Adkinson NF Jr, Bock SA, Branum A, Brown SG, Camargo CA Jr, Cydulka R, Galli SJ, Gidudu J, Gruchalla RS, Harlor AD Jr, Hepner DL, Lewis LM, Lieberman PL, Metcalfe DD, O'Connor R, Muraro A, Rudman A, Schmitt C, Scherrer D, Simons FE, Thomas S, Wood JP, Decker WW. Second symposium on the definition and management of anaphylaxis: Summary report-Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium. *J Allergy Clin Immunol*. 2006;117(2):391–397.

Saris-Baglama, RN, Dewey, CJ, Chisholm, GB, Kosinski, M, Bjorner, JB, & Ware Jr, JR (2004). SF health outcomes™ scoring software user's guide. Lincoln, RI: QualityMetric Inc. University of Pittsburgh. IDS/QIDS resource page. Available at: http://www.ids-qids.org/. Accessed April 2, 2015.

Stewart A, Hays R, Ware J. Methods of validating MOS health measures. In: Stewart A, Ware J, eds. Measuring Functioning and Well-being: The Medical Outcomes Study Approach. Durham, NC: Duke University Press; 1992:25-39.

Wyrwich KW, Nienaber NA, Tierney WM,Wolinksy FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care.* 1999a;37(5):469-478.

Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol*. 1999b;52(9):861-873.

Yates A. *Multivariate exploratory data analysis: A perspective on exploratory factor analysis.* Suny Press; 1987.

Zumbo BD. A handbook on the theory and methods of differential item functioning (DIF). *Ottawa: National Defense Headquarters.* 1999.

9. Appendices

Appendix 1. Derivation of SF-36v2® Health Survey, Acute Version Scores

The SF-36v2® Health Survey Scoring Software (QualityMetric Health Outcomes™ Scoring Software 4.5) will be used to calculate the SF-36v2® 8-domain and 2-component summary scores (Saris-Baglama et al. 2004). The SF-36 is a 36-item, patient-completed measure designed to be a short, multipurpose assessment of health in the areas of physical functioning, role-physical, role-emotional, bodily pain, vitality, social functioning, mental health, and general health. The 2 overarching dimensions of mental well-being and physical well-being are captured by the Physical Component Summary (PCS) and Mental Component Summary (MCS) scores, respectively. The summary scores range from 0 to 100, with higher scores indicating better levels of function and/or better health. Items are answered on Likert scales of varying lengths. In this study, the SF-36 acute version will be used, which has a 1-week recall period (The SF Community – SF-36 Health Survey Update [WWW]).

The Scoring Software performs a 4-step process to calculate raw domain scores and t-scores consisting of

1. Data Cleaning and Item recording: First, data are checked for out of range values, converting invalid items to missing values. Next, items (BP01, BP02, GH01, GH03, GH05, VT01, VT02, SF01, MH03, MH05) are reverse scored, so that higher scores denote better health for all SF-36 items.
2. Although current research indicates a linear relationship between the SF-36 items and the underlying health concept, empirical research suggests that items GH01 and BP01 require recalibration to satisfy important scaling assumptions. Item GH01 will be rescored according to the following table:

Response to GH01	Recommended Value / Recalibrated Value
Excellent	5.0
Very Good	4.4
Good	3.4
Fair	2.0
Poor	1.0

The BP01 will be rescored with

Response Choices	Final Item Value
None	6.0
Very mild	5.4
Mild	4.2
Moderate	3.1
Severe	2.2
Very severe	1.0

Item 08 (BP02) will be rescored if BP01 and BP02 were answered

Response Choices	If BP02 Pre-coded Item Value	and BP01 Pre-coded Item Value	Then Final Item 08 (BP02) Value
Not at all	1	1	6
Not at all	1	2-6	5
A little bit	2	1-6	4
Moderately	3	1-6	3
Quite a bit	4	1-6	2
Extremely	5	1-6	1

Scoring of BP02 if BP01 is not answered:

Response Choices	Final Item Value
Not at all	6.0
A little bit	4.75
Moderately	3.5
Quite a bit	2.25
Extremely	1.0

3. After this rescoring, the raw domain scores will be calculated for the scale. Domain scores are the simple algebraic sum of the final values for all items in that scale.
4. All raw domain scores will be transformed to a 0-100 scale, with 0 being the lowest and 100 the highest possible score.
5. Finally, the 0-100 scores will be transformed to t-score based scores. First, a z-score transformation using the mean for the respective recall period, here 1 week recall, of the 1998 general U.S. population will be used. Then the distribution of z-score is linearly transformed to have a mean of 50 and a SD of 10 by multiplying each z-score with 10 and adding 50.

The calculation of component scores is a 3-step process using the domain scores, calculated as described above:

1. The standardized scores from Step 5, depending on the chosen recall period, are calculated.
2. These standardized Physical and Mental component scores are calculated as the weighted sums by the factor score coefficients, derived from the 1990 general U.S. population, with the domain scores. If any domain score is missing then the aggregate Physical or Mental score will not be calculated.
3. The PCS and MCS are linearly transformed by multiplying by 10 and adding 50 to obtain the aggregate t-score based scoring.

To run the scoring algorithm, the SF-36 items recorded in the study database will be exported into a comma- or tab-separated values file (*.csv, *.tab). This file will then be loaded into the Scoring Software to perform the calculations described above. The resulting raw domain scores and t-scores (domain scores) will then be exported into a comma- or tab-separated values file and imported into SAS for storage in the SDTM/ADaM datasets.

The comma- or tab-separated values file, each row will be one patient record and the first row will comprise the header columns, will have the following column specification: (to comply with the Scoring Software requirements)

eCRF Row #	**Column label for export to comma- or tab-separated values file [*.csv, *.tab]**	**Annotated SF-36 eCRF Variable** [Format: SF36V2RXX_SF36V2F1**]**	**Item Number, Score range**	**eCRF question / Specification**
1	GH01	[SF36V2R01_SF36V2F1]	Item # 1, Range 1-5	In general, would you say your health is:
2	HT	[SF36V2R02_SF36V2F1]	Item # 2, Range 1-5	Compared to one week ago, how would you rate your health in general now?
3	PF01	[SF36V2R03_SF36V2F1]	Item # 3a, Range 1-3	Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports
4	PF02	[SF36V2R04_SF36V2F1]	Item # 3b, Range 1-3	Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
5	PF03	[SF36V2R05_SF36V2F1]	Item # 3c, Range 1-3	Lifting or carrying groceries
6	PF04	[SF36V2R06_SF36V2F1]	Item # 3d, Range 1-3	Climbing several flights of stairs
7	PF05	[SF36V2R07_SF36V2F1]	Item # 3e, Range 1-3	Climbing one flight of stairs
8	PF06	[SF36V2R08_SF36V2F1]	Item # 3f, Range 1-3	Bending, kneeling, or stooping
9	PF07	[SF36V2R09_SF36V2F1]	Item # 3g, Range 1-3	Walking more than a mile
10	PF08	[SF36V2R10_SF36V2F1]	Item # 3h, Range 1-3	Walking several hundred yards
11	PF09	[SF36V2R11_SF36V2F1]	Item # 3i, Range 1-3	Walking one hundred yards
12	PF10	[SF36V2R12_SF36V2F1]	Item # 3j, Range 1-3	Bathing or dressing yourself
13	RP01	[SF36V2R13_SF36V2F1]	Item # 4a, Range 1-5	Cut down the amount of time you spent on work or other activities
14	RP02	[SF36V2R14_SF36V2F1]	Item # 4b, Range 1-5	Accomplished less than you would like
15	RP03	[SF36V2R15_SF36V2F1]	Item # 4c, Range 1-5	Were limited in the kind of work or other activities

eCRF Row #	Column label for export to comma- or tab-separated values file [*.csv, *.tab]	Annotated SF-36 eCRF Variable [Format: SF36V2RXX_SF36V2F1]	Item Number, Score range	eCRF question / Specification
16	RP04	[SF36V2R16_SF36V2F1]	Item # 4d, Range 1-5	Had difficulty performing the work or other activities (for example, it took extra effort)
17	RE01	[SF36V2R17_SF36V2F1]	Item # 5a, Range 1-5	Cut down the amount of time you spent on work or other activities
18	RE02	[SF36V2R18_SF36V2F1]	Item # 5b, Range 1-5	Accomplished less than you would like
19	RE03	[SF36V2R19_SF36V2F1]	Item # 5c, Range 1-5	Did work or other activities less carefully than usual
20	SF01	[SF36V2R20_SF36V2F1]	Item # 6, Range 1-5	During the past week, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?
21	BP01	[SF36V2R21_SF36V2F1]	Item # 7, Range 1-6	How much bodily pain have you had during the past week?
22	BP02	[SF36V2R22_SF36V2F1]	Item # 8, Range 1-5	During the past week, how much did pain interfere with your normal work (including both work outside the home and housework)?
23	VT01	[SF36V2R23_SF36V2F1]	Item # 9a, Range 1-5	Did you feel full of life?
24	MH01	[SF36V2R24_SF36V2F1]	Item # 9b, Range 1-5	Have you been very nervous?
25	MH02	[SF36V2R25_SF36V2F1]	Item # 9c, Range 1-5	Have you felt so down in the dumps that nothing could cheer you up?
26	MH03	[SF36V2R26_SF36V2F1]	Item # 9d, Range 1-5	Have you felt calm and peaceful?
27	VT02	[SF36V2R27_SF36V2F1]	Item # 9e, Range 1-5	Did you have a lot of energy?
28	MH04	[SF36V2R28_SF36V2F1]	Item # 9f, Range 1-5	Have you felt downhearted and depressed?
29	VT03	[SF36V2R29_SF36V2F1]	Item # 9g, Range 1-5	Did you feel worn out?
30	MH05	[SF36V2R30_SF36V2F1]	Item # 9h, Range 1-5	Have you been happy?
31	VT04	[SF36V2R31_SF36V2F1]	Item # 9i, Range 1-5	Did you feel tired?
32	SF02	[SF36V2R32_SF36V2F1]	Item # 10, Range 1-5	During the past week, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

eCRF Row #	Column label for export to comma- or tab-separated values file [*.csv, *.tab]	Annotated SF-36 eCRF Variable [Format: SF36V2RXX_SF36V2F1]	Item Number, Score range	eCRF question / Specification
33	GH02	[SF36V2R33_SF36V2F1]	Item # 11a, Range 1-5	I seem to get sick a little easier than other people
34	GH03	[SF36V2R34_SF36V2F1]	Item # 11b, Range 1-5	I am as healthy as anybody I know
35	GH04	[SF36V2R35_SF36V2F1]	Item # 11c, Range 1-5	I expect my health to get worse
36	GH05	[SF36V2R36_SF36V2F1]	Item # 11d, Range 1-5	My health is excellent
	SUBJID			Subjects ID
	VISID			Visid ID
	Sex			Sex coded as: f/m
	DateOfBirth			Date of birth formatted as: mm/dd/yyyy (when scoring software is run in US), or dd/mm/yyyy (when scoring software is run in non-US)
	RecordID			Running number for the exported records

The SF-36 Scoring Software will derive raw domain scores and t-scores that can be exported into a comma- or tab-separated values file (*.csv, *.tab) with the following columns added:

Column label added to comma- or tab-separated values file from export [*.csv, *.tab]	Scoring Software specification
PF	Physical Functioning domain score
RP	Role Limitations Due To Physical Health domain score
BP	Bodily Pain domain score
GH	General Health Perceptions domain score
VT	Vitality domain score
SF	Social Functioning domain score
RE	Role Limitations Due To Emotional Problems domain score
MH	Mental Health domain score
PCS	Physical Component score
MCS	Mental health Component score

For scoring the trial data, the Missing Data Estimator option will not be selected. If an item is missing, there will be no imputation conducted by the Scoring Software. Only complete questionnaire data will be scored.

The Scoring Software also allows for calculating domain and component scores from weights derived from an oblique factor solution for comparative purposes. This option will not be used.

The SF-36 scoring using the Scoring Software will be conducted by a 2 person team overseeing each other in a single scoring session. In case of relevant observations during the scoring, those will be documented in pertinent meeting minutes and filed as part of the study documentation.