Janssen Research & Development

Statistical Analysis Plan

A Phase 3B, Multicenter, Randomized, Double-blind, Placebo-controlled Study Evaluating the Efficacy and Safety of Guselkumab Administered Subcutaneously in Participants with Active Psoriatic Arthritis who had an Inadequate Response and/or Intolerance to One Prior Anti-Tumor Necrosis Factor α Agent

Protocol CNTO1959PSA3005; Phase 3b
Amendment 1
CNTO1959 (guselkumab)

Status:	Approved
Date:	20 November 2024
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-RIM-1239659, 2.0

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

LIST OF IN TEXT TABLES AND FIGURES

Tables

Figures

VERSION HISTORY

Table 1: SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	06 February 2024	Not Applicable	Initial release
2		• Added clinically important hepatic disorder event summary table and listing; added MACE listing; added opportunistic infections listing • Removed language involving possible Hy's Law, and replaced it with 2 combined Biochemical criteria in Section 5.4.4.1.	For consistency with other studies for the compound
		• Added SPARCC based enthesitis definition • In Dactylitis related Other Endpoints, added 'Proportion of participants who achieve resolution of dactylitis by visit over time through Week 100 in those participants with dactylitis at baseline' • In DAPSA related Other Endpoints, removed 'Participants in remission, or with low disease activity based on the DAPSA score are those participants who have a DAPSA score less than or equal to 4, or less than or equal to 14, respectively' • Defined the categories in the Participating Regions subgroup • In Appendix 7, added a table on how AEs of interest are categorized • Various typographical or internal consistency corrections	Other

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for the analysis of efficacy, safety, pharmacokinetics (PK), and pharmacodynamics (PD), and Immunogenicity in the CNTO1959PSA3005 study.

1.1. Objectives and Endpoints

Primary Objective

The primary objective of this study is to evaluate the efficacy of guselkumab treatment in participants with active psoriatic arthritis (PsA) and inadequate response (IR) and/or intolerance to one prior anti-TNF by assessing the reduction in signs and symptoms of PsA. The proportion of participants who achieve an ACR 20 response at Week 24 will be used for this assessment.

Major Secondary Objective

The secondary objective of this study is to evaluate the efficacy of guselkumab on additional measures of signs and symptoms of PsA, psoriasis, and patient well-being. These objectives will be assessed by:

1. Proportion of participants who achieve a psoriasis response of Investigator's Global Assessment (IGA) psoriasis score of 0 (cleared) or 1 (minimal) AND ≥2-grade reduction from baseline at Week 24 among the participants with ≥3% body surface area (BSA) psoriatic involvement and an IGA score of ≥2 (mild) at baseline
2. Proportion of participants who achieve Psoriasis Area and Severity Index (PASI) 90 response at Week 24 among the participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline
3. Change from baseline in HAQ disability index (HAQ-DI) at Week 24
4. Change from baseline in the 36-Item Short Form Survey Instrument (SF-36) Physical Component Summary (PCS) at Week 24
5. Change from baseline in Functional Assessment of Chronic Illness Therapy - fatigue (FACIT-F) score at Week 24
6. Proportion of participants achieving minimal disease activity (MDA) at Week 24
7. Proportion of participants who achieve ACR20 response at Week 16
8. Proportion of participants who achieve ACR50 response at Week 16
9. Proportion of participants who achieve ACR50 response at Week 24
10. Proportion of participants who achieve ACR70 response at Week 24

Other Secondary Objectives

Other secondary objectives of this study are:

1. To evaluate the safety of guselkumab in participants with active PsA. These objectives will be assessed by:

a. Frequency and type of adverse events (AEs), serious AEs (SAEs), reasonably related AEs, AEs leading to discontinuation of study intervention, infections, infusion reactions, and injection-site reactions.

b. Laboratory abnormalities (chemistry, hematology), maximum toxicity (Common Terminology Criteria for Adverse Events [CTCAE 5.0]) grades.

2. To evaluate the PK and immunogenicity of guselkumab in participants with active PsA. These objectives will be assessed by:

   a. Serum guselkumab concentration over time

   b. Incidence of antibodies to guselkumab

Other Endpoints:

CCI

CCI

[redacted]

1.2. Study Design

This is a Phase 3b, multicenter, randomized, double-blind, placebo-controlled interventional study in participants with active PsA who had an inadequate response and/or intolerance to one prior anti-TNF.

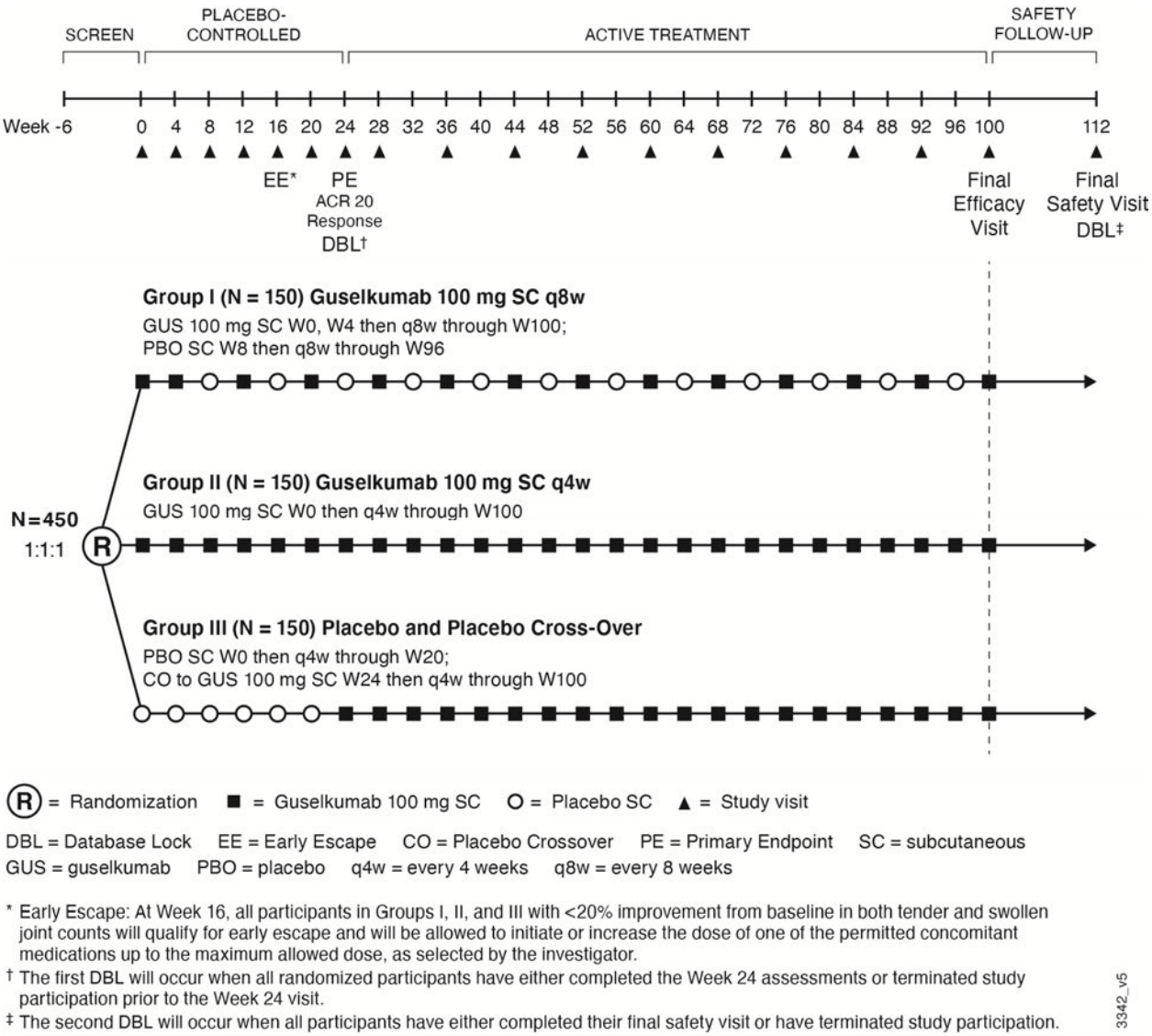
A target of 450 participants will be randomly assigned in this study with 150 participants planned per intervention group: guselkumab q8w (n=150), guselkumab q4w (n=150), or placebo (n=150).

Stable doses of concomitant non-steroidal anti-inflammatory drugs (NSAIDs), oral corticosteroids (≤10 mg/day prednisone equivalent) and selected non-biologic DMARDs (methotrexate [MTX], sulfasalazine [SSZ], hydroxychloroquine [HCQ], leflunomide [LEF]) will be allowed but are not required.

There will be a screening phase of approximately 6 weeks, a treatment phase of approximately 2 years that will include a double-blind placebo-controlled period from Week 0 to Week 24 and an active controlled treatment phase from Week 24 to Week 100, and a safety follow-up at Week 112 (approximately 12 weeks after the last intended dose at Week 100 per protocol).

An overview of the study design is provided in Figure 1.

Figure 1: Schematic Overview of the Study Through End of Study



At Week 0, participants who satisfy all inclusion and exclusion criteria will be randomly assigned to 1 of the following 3 treatment groups in a 1:1:1 ratio using permuted block randomization stratified by baseline non-biologic DMARD use (yes/no).

- **Group I (n=150), guselkumab 100 mg SC every 8 weeks (q8w):** Participants will receive SC guselkumab 100 mg at Weeks 0, 4 then q8 at Weeks 12, 20, 28, 36, 44, 52, 60, 68, 76, 84, 92, and 100. Participants will receive placebo SC at Weeks 8, 16, 24, 32, 40, 48, 56, 64, 72, 80, 88, and 96 to maintain the blind.
- **Group II (n = 150), guselkumab 100 mg SC every 4 weeks (q4w):** Participants will receive SC guselkumab 100 mg at Weeks 0, 4, then q4 at Weeks 8, 12, 16, 20, 24, 28, 32, 36, 40, 44 48, 52, 56, 60, 64, 68, 72, 76, 80, 84, 88, 92, 96, and 100.
- **Group III (n=150), placebo and placebo cross-over:** Participants will receive SC placebo at Weeks 0, 4, 8, 12, 16 and 20, and will cross over at Week 24 to receive SC guselkumab

100 mg at Weeks 24, 28, then q4w at Weeks 32, 36, 40, 44, 48, 52, 56, 60, 64, 68, 72, 76, 80, 84, 88, 92, 96, and 100.

At Week 16, all participants in Groups I, II, and III with <20% improvement from baseline in both tender and swollen joint counts will qualify for early escape (EE) and will be allowed to initiate or increase the dose of one of the permitted concomitant medications up to the maximum allowed dose as specified in Table 2, as selected by the investigator.

Table 2: Permitted Concomitant Medications for PsA and the Maximum Allowed Doses During the Study

Permitted Concomitant Medications for PsA[a]	**Maximum Allowed Dose**
NSAIDs and other analgesics	Marketed dose approved in in the country where the study is being conducted
Oral corticosteroids	Equivalent to 10 mg/day of prednisone
Methotrexate (MTX)[b]	25 mg/week
Sulfasalazine (SSZ)	3 g/day
Hydroxychloroquine (HCQ)	400 mg/day
Leflunomide (LEF)	20 mg/day

[a] Permitted concomitant medication are not supplied by the Sponsor.
[b] It is recommended that all participants taking MTX in this study receive at least 5 mg oral folate or 5 mg folinic acid weekly. Guidelines for dose adjustment in the event of MTX toxicity are included in the Trial Center File.

At Week 24, all participants in the placebo group (Group III) will cross over in a blinded fashion to receive guselkumab 100 mg SC q4w through Week 100; this will allow for collection of additional safety data for the guselkumab 100 mg SC q4w dose. Participants in the guselkumab groups (Groups I and II) will remain on the dosing regimen they were randomized to at Week 0, through Week 100.

Database locks (DBLs) are scheduled at Weeks 24, Week 52, and the end of the study (Week 112). The first DBL will occur when all randomized participants have either completed the Week 24 assessments or terminated study participation prior to the Week 24 visit (referred to as Week 24 DBL). The second DBL will occur when all randomized participants have either completed the Week 52 assessments or terminated study participation prior to the Week 52 visit (referred to as Week 52 DBL). The third DBL will occur when all participants have either completed their final safety visit or have terminated study participation (referred to as Final DBL).

The end of study is considered as the last visit for the last participant in the study.

Blinding

To maintain the study blind, the study intervention container will have a label containing the study name, study intervention number, reference number and storage instructions. The label will not identify the study intervention in the container. However, if it is necessary for a participant's safety, the study blind may be broken, and the identity of the study intervention ascertained. The study

intervention number will be entered in the electronic case report form (eCRF) when the study intervention is administered. The study interventions will be identical in appearance and will be packaged in identical containers.

The investigator will not be provided with randomization codes. The codes will be maintained within the interactive web response system (IWRS), which has the functionality to allow the investigator to break the blind for an individual participant.

Data that may potentially unblind the intervention assignment (ie, study intervention serum concentrations, anti-guselkumab antibodies levels) will be handled with special care to ensure that the integrity of the blind is maintained and the potential for bias is minimized. This can include making special provisions, such as segregating the data in question from view by the investigators, clinical team, or others as appropriate until the time of database lock and unblinding.

The results of the post-baseline C-reactive protein (CRP) measurements performed by the central laboratory will be blinded to the investigative sites.

Under normal circumstances, the blind should not be broken until all participants have completed the study and the database has been finalized. The investigator may in an emergency determine the identity of the intervention by contacting the IWRS. While the responsibility to break the intervention code in emergency situations resides solely with the investigator, it is recommended that the investigator contacts the Sponsor or its designee, if possible, to discuss the particular situation, before breaking the blind. Telephone contact with the Sponsor or its designee will be available 24 hours per day, 7 days per week. In the event the blind is broken, the Sponsor must be informed as soon as possible. The date and reason for the unblinding must be documented by the IWRS, in the appropriate section of the eCRF, and in the source document. The documentation received from the IWRS indicating the code break must be retained with the participant's source documents in a secure manner.

Participants who have had their intervention assignment unblinded should continue to return for scheduled evaluations and may not be eligible for further treatment.

At the Week 24 DBL and Week 52 DBL, the data will be unblinded for analysis to some Sponsor personnel while participants are still participating in the study. Identification of Sponsor personnel who will have access to the unblinded subject-level data will be documented prior to unblinding. Investigative study sites and participants will remain blinded to initial treatment assignment until after the final database is locked and analyses completed.

2. STATISTICAL HYPOTHESES

The primary efficacy endpoint of this study is the proportion of participants who achieved an ACR 20 response at Week 24 (refer to Section 5.3.4 for endpoint definition and analyses). This endpoint was chosen because it is well-accepted by regulatory authorities and the clinical PsA community.

The hypotheses related to the primary endpoint are that:

1. **(H1)** Treatment with guselkumab 100 mg SC q4w is superior to treatment with placebo SC with respect to reduction of PsA signs and symptoms as measured by proportion of participants who achieved an ACR 20 response at Week 24 (***primary hypothesis***); and
2. **(H2)** Treatment with guselkumab 100 mg SC at Week 0, Week 4 and then q8w is superior to treatment with placebo SC with respect to reduction of PsA signs and symptoms as measured by proportion of participants who achieved an ACR 20 response at Week 24 (key ***secondary hypotheses*).**

The first hypothesis (H1) is the **primary hypothesis** for this study. If the first hypothesis achieves the statistical significance at a 2-sided α-level of 0.05, the study will be considered positive.

In addition to the primary endpoint, there are 10 major secondary endpoints in this study (refer to Section 5.3.5 for endpoint definitions and analyses). The hypotheses related to the major secondary (all are major secondary hypotheses) are provided in Section 6.10 (Appendix 10).

For hypothesis testing order and multiplicity adjustment, refer to Section 5.3.3.2.

3. SAMPLE SIZE DETERMINATION

The planned enrollment in the study is approximately 450 participants. The sample size selection was determined based on the primary endpoint of proportion of participants who achieve an ACR 20 response at Week 24. The assumptions are based on the PSA3001 and PSA3003 studies.

3.1. Primary Endpoint – ACR 20 Response at Week 24

In the PSA3001 study, the ACR 20 response rates in the overall efficacy population at Week 24 were 22.2%, 52.0%, and 59.4%, for the placebo, guselkumab SC 100 mg q8w, and guselkumab 100 mg q4w treatment groups, respectively.

In the PSA3003 study which included participants who have an inadequate response or intolerance to up to 2 prior anti-TNFs, the ACR 20 response rates at Week 24 were 19.8% and 44.4%, for placebo and guselkumab SC 100 mg q8w, respectively.

For this study, assuming an ACR 20 response rate comparable to those in PSA3001 and PSA3003 in each of the guselkumab groups and the placebo group, a sample size of 150 participants per group for each of the guselkumab treatment groups and placebo will have >90% power to detect differences between each guselkumab treatment group and placebo for the primary endpoint assuming a 2-sided alpha level of 0.05 (Table 3). The sample size calculations assume that 100% of the effect observed in PSA3001 and PSA3003 will be observed. If only 70 or 80% of the effect seen in PSA3001 and PSA3003 is observed, the power will still exceed 90% for the Primary endpoint.

Table 3: Power Calculations for Primary Endpoint Based on 150 per Treatment Group

Endpoint	Source		Mean or proportion		Difference (Δ)	Power (n=150 p=0.05)
			Placebo Group	Guselkumab Group		
Primary: ACR20 at Week 24						
	PSA3001	Q8	22%	52%	30%	>99%
	PSA3001	Q4	22%	59%	37%	>99%
	PSA3003[a]	Q8	19.8%	44.4%	24.6%	>99%
[a:] PSA3003 (COSMOS) data is for Q8W only						

3.2. Major Secondary

Table 4 provides the statistical power for major secondary with 150 participants per treatment group. The study power ranges approximately from 44% to 99%.

Table 4: Power Calculations for Major Secondary Endpoint Based on 150 Participants per Treatment Group

Endpoint	Source		Mean or proportion				
			Placebo	Guselkumab Group	SD[a]	Difference (Δ)	Power (n=150 p=0.05)
(multiplicity controlled)							
IGA at Week 24	PSA3001	Q8	15%	57%		42%	>99%
	PSA3001	Q4	15%	75%		60%	>99%
PASI90	PSA3001	Q8	11.5%	50.0%		38.5%	>99%
	PSA3001	Q4	11.5%	62.9%		51.4%	>99%
	PSA3003[b]	Q8	7.5%	51.1%		43.6%	>99%
HAQ-DI at Week 24	PSA3001	Q8	-0.07	-0.32	0.56	0.25	97%
	PSA3001	Q4	-0.07	-0.40	0.56	0.33	>99%
	PSA3003[b]	Q8	-0.009	-0.178	0.52	0.169	>99%
SF-36 (PCS) at Week 24	PSA3001	Q8	1.96	6.10	7.72	4.14	>99%
	PSA3001	Q4	1.96	6.87	7.72	4.91	>99%
	PSA3003[b]	Q8	-0.387	3.514	6.61	3.901	>99%
FACIT-F at Week 24	PSA3002	Q8	3.56	7.55	9.87	3.99	87%
	PSA3002	Q4	3.56	7.11	9.87	3.55	93%
	PSA3003[b]	Q8	1.050	4.607	8.67	3.557	94%
MDA	PSA3001	Q8	11.1%	22.8%		11.7%	78%
	PSA3001	Q4	11.1%	30.5%		19.4%	99%
	PSA3003[b]	Q8	3.1%	14.8%		11.7%	95%
Major secondary (weakly controlled)							
ACR20 at Week 16	PSA3001	Q8	25%	52%		27%	>99%
	PSA3001	Q4	25%	60%		35%	>99%
	PSA3003[b]	Q8	16.7%	41.3%		24.6%	>99%

Endpoint	Source		Mean or proportion				
			Placebo	Guselkumab Group	SD[a]	Difference (Δ)	Power (n=150 p=0.05)
ACR50 at Week 24	PSA3001	Q8	9%	30%		21%	>99
	PSA3001	Q4	9%	36%		27%	>99%
	PSA3003[b]	Q8	5.2%	19.6%		14.4%	97%
ACR50 at Week 16	PSA3001	Q8	13%	23%		10%	62%
	PSA3001	Q4	13%	27%		14%	87%
	PSA3003[b]	Q8	5.2%	14.3%		9.1%	77%
ACR70 at Week 24	PSA3001	Q8	6%	12%		6%	44%
	PSA3001	Q4	6%	20%		14%	96%
	PSA3003[b]	Q8	1.0%	7.9%		6.9%	84%

[a]: Used the largest standard deviation among the treatment groups.
[b]: PSA3003 (COSMOS) data is for Q8W only.

4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS

4.1. Enrolled Participants

All participants who signed the informed consent form (ICF).

4.2. Full Analysis Set (FAS) (Week 0 – Week 24)

All participants who were randomized in the study. This analysis set will be used for efficacy analyses.

This SAP uses the following definitions:

- Natural Disaster includes the COVID-19 pandemic, specifically any site closures, site access restrictions, or lockdowns caused by the pandemic;
- Major Disruption includes the disruption in Ukraine and neighboring countries/territories beginning 24 February 2022.

Because there were no sites where **every** participant was unable to have data collected for the primary endpoint at Week 24 or missed ≥2 doses of study intervention prior to Week 24, all the sites will be included in the FAS.

In the analyses for this set, participants will be analyzed according to the randomized study intervention they were **assigned to**, regardless of the study intervention they actually received.

4.3. Full Analysis Set Excluding Ukraine (FAS-UKR) (Week 0 – Week 24)

All participants in FAS except those from study sites in Ukraine will be included in the FAS-UKR analysis set. This analysis set will be used for efficacy analyses.

FAS-UKR is being utilized as no participant from Ukrainian sites reached Week 24 prior to 24 February 2022 and Ukrainian sites were potentially impacted in various ways: study intervention supply interruptions (supply depot to the entire country cut off for a period of time),

temporary site closures in some sites, and other difficulties involved in operation within an area of ongoing conflict. Thus, while all sites will be included in FAS, this analysis set will be used for sensitivity and supplementary analyses of the Primary endpoints, as well as supportive analyses for the Major Secondary and Other Efficacy endpoints based on the Treatment Policy Estimand.

In the analyses for this set, participants will be analyzed according to the randomized study intervention they were **assigned to**, regardless of the study intervention they actually received.

4.4. Per-Protocol Analysis Set (Week 0 – Week 24)

The Per-Protocol Analysis Set (PPAS) includes all participants in FAS who met all inclusion and exclusion criteria and had no major protocol deviations prior to Week 24 that could have impacted efficacy assessment per clinical judgement (Appendix 3). This analysis set will be used for the analyses of ***selected*** efficacy endpoints through Week 24. Participants to be excluded from this analysis will be identified prior to the Week-24 DBL and unblinding.

In the analyses for this set, participants will be analyzed according to the randomized study intervention they were **assigned to**, regardless of the study intervention they actually received.

4.5. Safety Analysis Set

All participants who received at least one (complete or partial) administration of any study intervention, ie, the treated population. This analysis set will be used for the safety analyses.

In the analyses for this set, participants will be analyzed per the study intervention they **actually received**, regardless of the study intervention they were randomized to.

4.6. Pharmacokinetics (PK) Analysis Set

All participants who received at least one complete administration of guselkumab and had at least one valid blood sample drawn for PK analysis.

4.7. Immunogenicity Analysis Set

All participants who received at least one (complete or partial) administration of guselkumab and who had at least 1 sample obtained after their first administration of guselkumab.

4.8. Pharmacodynamic (PD) Analysis Set

The PD analysis set will be defined in a separate analysis plan for biomarkers and PD analyses.

5. STATISTICAL ANALYSES

5.1. General Considerations

5.1.1. Visit Windows

In general, for safety and efficacy analyses, the nominal visit will be used for by-visit analyses. Exceptions exist for the following.

5.1.1.1. Visit Windows for Dosing and PK Analysis

All post-baseline visits from baseline through Week 24 will have a visit window of ±4 days, and from Week 28 through Week 100 will have a visit window of ±7 days. Week 112 or the final safety visit (12 weeks after the last dose in participants who discontinued study treatment) will have a visit window of ±14 days. This information will be used to identify out-of-window dosing or visits.

For PK analyses, if a participant has an administration outside the visit window at a visit, the concentration data collected at and after that visit will be excluded from the by-visit data analyses.

5.1.2. Pooling Algorithm for Analysis

Data from all investigational centers/sites will be pooled for analyses.

5.2. Participant Dispositions

The number of participants screened, randomized and treated will be summarized by geographic region, country, and investigational site.

The number of participants in the following disposition categories will be summarized throughout the study by intervention group and overall.

- Participants who received study intervention
- Participants who completed the study
- Participants who discontinued study intervention
- Reasons for discontinuation of study intervention
- Participants who terminated study prematurely
- Reasons for termination of study

Listings of participants will be provided for the following categories:

- Participants who discontinued study intervention
- Participants who terminated study prematurely
- Participants who were unblinded prior to the Week 112 DBL
- Participants who were randomized yet did not receive study intervention
- Summaries of participant demographic and baseline characteristics are in Section 6.2 (Appendix 2).

5.2.1. Intercurrent Events (ICEs) and Early Escape (EE)

Tabulations by randomized intervention group will be provided for participants who met EE criteria at Week 16 and for participants who met 1 or more ICEs for the Adjusted Composite Estimand prior to Week 24 defined in Section 5.3.4.2.1.

In addition, participants who met any ICE criteria will be presented in data listings.

5.3. Efficacy Analyses

5.3.1. General Method of Analysis

In general, descriptive statistics, such as mean, standard deviation (SD), median, interquartile (IQ) range, minimum, and maximum for continuous variables, and counts and percentages for discrete variables will be used to summarize most data.

Statistical comparison between a guselkumab group (100mg q4w or 100mg at Weeks 0, 4 and then q8w) and the placebo group will be performed by visit through Week 24. No treatment comparisons will be performed after Week 24.

Binary Response Efficacy Endpoints

For binary response efficacy endpoints where any portion of the missing data is imputed using Multiple Imputation (MI) (see Section 6.12 for technical details), treatment comparisons will be performed using a **Cochran-Mantel-Haenszel (CMH)** stratified by baseline use of non-biologic DMARDs (yes, no) on each of the imputation sets, and the inferences from the analysis of each imputed data are pooled. The within imputation variance and between imputation variance are combined to estimate the total variance of the stratum adjusted difference of proportions. This estimate of the variance and critical values from the t-distribution are used to calculate the confidence interval for the stratum adjusted difference of proportions. The SAS procedure PROC MIANALYZE is used where the critical value is based on the t-distribution which is different from the analysis not based on MI where normal distribution is used. The large number of observations in our data imply that the critical values from the t-distribution are almost identical to the critical values from the standard normal distribution. The Wilson-Hilferty (Ratitch 2013) transformation is used to pool the CMH statistics from each imputed data set to calculate the p-value.

For analyses where all missing data is imputed only a single time (eg, non-responder imputation for all missings) or not imputed with the missing assumed Missing Completely at Random (MCAR), treatment comparisons will generally be performed using a CMH test stratified by baseline use of non-biologic DMARDs (Yes/No). The magnitude of the treatment difference will be estimated by the difference in response rates between the guselkumab and placebo groups with a 95% confidence interval (CI) calculated based on Wald statistics (Kim 2013). The Mantel Fleiss criterion will be used to determine the appropriateness of using the CMH test. If the Mantel Fleiss criterion is not satisfied, the Fisher's exact test will be used instead of the CMH test to compare the two intervention groups.

For endpoints where any portion of the missing data is not imputed, but still needs to be accounted for under Missing at Random (MAR) assumptions, the **Generalized Linear Mixed Model (GLMM)** will be used (for more details see Appendix 13 in Section 6.13).

For subgroup analyses, a **logistic regression model** using data from all 3 intervention groups will be used to obtain odds ratios and associated 95% CI. If any missing data were multiply imputed, the logistic regression model will be run on each of the imputation sets, and the odds ratios from each imputation set will have the log transformation applied before pooling across imputation sets.

Once combined, the log transformed odds ratio and 95% CI will be back transformed to their original scale (Ratitch 2013) and presented. The treatment*subgroup interaction p-value will be obtained separately, based on logistic regression on the observed data and with only the data from the 2 intervention groups currently being analyzed fed into the model.

Continuous Efficacy Endpoints

For most continuous endpoints where any portion of the missing data is imputed using MI (see Appendix 12 in Section 6.12 for technical details), treatment comparisons will be performed using an **Analysis of Covariance (ANCOVA)** model based on each of the imputation sets. The estimate of the mean change from baseline is the average of the mean change taken over all the MI data sets. The estimate of the variance of the mean change from baseline is the weighted sum of the average within-imputation variance and the between-imputation variance. The CI for the mean change from baseline uses critical values from the t-distribution. The treatment difference between each guselkumab group versus the placebo group will be tested for each imputation dataset and then the analysis results across all imputation datasets will be combined. The treatment difference in the change from baseline is estimated by the average of the treatment differences over the MI data sets. The estimate of the variance of the treatment difference in the change from baseline is the weighted sum of the average within-imputation variance and the between-imputation variance, under the assumption of homogeneity of variance between intervention groups for performing ANCOVA within each imputation dataset. The CI is based on the critical values from the t-distribution. The large number of observations in our data imply that the critical values from the t-distribution are almost identical to the critical values from the standard normal distribution.

The ANCOVA model will be based on the original scale and will include intervention group, baseline score, and baseline use of non-biologic DMARDs (yes, no) as the explanatory factors. The model will include data from all the 3 intervention groups.

For analyses where all missing data is imputed only a single time (eg, change=0 for all missings) or not imputed with the missing assumed MCAR, treatment comparisons will generally be performed using ANCOVA.

For endpoints where any portion of the missing data is not imputed, but still needs to be accounted for under MAR assumptions, the **Mixed-Effect Model Repeated Measures (MMRM)** model will be used (for more details see Appendix 13 in Section 6.13).

5.3.2. General Data Handling Rules

Missing data will be handled depending on the assumed mechanism behind the missingness, whether the endpoint is a key endpoint, and the estimand being used. Handling rules are summarized in Table 5 below:

Table 5: Data Handling Rules for Missing Data

Estimand	Assumption	Continuous endpoints	Binary response endpoints
Adjusted Composite	MAR (Missing at Random) due to Natural Disaster or Major Disruption	**Major secondary endpoint**: imputed using Full Conditional Specifications (FCS) MI **Clinical Other endpoints**: not imputed but accounted for in MMRM model	**Primary and Major secondary endpoints**: imputed using Full Conditional Specifications (FCS) MI on continuous component scale then dichotomized **Clinical Other endpoints**: not imputed but accounted for in GLMM
	MAR NOT due to above reasons	**Major secondary endpoints:** imputed using Full Conditional Specifications (FCS) MI **Clinical Other endpoints:** not imputed but accounted for in MMRM model	**Primary and Major secondary endpoints:** conservatively imputed as non-response for consistency with historical studies **Clinical Other endpoints:** conservatively imputed as non-response for consistency with historical studies
	Missing Not at Random (MNAR)	N/A	**Primary endpoints:** ***For tipping point analyses specifically:*** Sensitivity analysis with exhaustive assessment of all possible combinations of response status for missing data
Treatment Policy	MAR	**Major secondary endpoints**: imputed using FCS MI	**Primary and Major secondary endpoints**: imputed using FCS MI on continuous component scale then dichotomized
	MNAR	**Major secondary endpoints:** ***For tipping point analyses specifically:*** Supplementary analysis systematically assessing scenarios which deviate from MAR or MCAR assumption	**Primary and Major secondary endpoints:** ***For tipping point analyses specifically:*** Supplementary analysis systematically assessing scenarios which deviate from MAR or MCAR assumption
	MCAR	Other Efficacy endpoints: not imputed	Other Efficacy endpoints: not imputed

5.3.2.1. MI Using FCS Regression

Under the assumption of MAR, MI will be used to impute the missing data for the continuous/ordinal measurements. The missing data will be imputed using the predicted value from an imputation model using the Full Conditional Specification (FCS) regression method for any missing pattern. Each variable will be restricted to only impute within its possible range of values (eg, HAQ Score may only be imputed to within 0-3). The explanatory variables in the imputation model include imputation variables and ancillary variables as specified in Section 6.12 (Appendix 12).

The number of imputations (N) and the starting seeds are also specified in Section 6.12 (Appendix 12).

For the composite continuous endpoints, the above imputation will be performed on components with missing data and then composite score will be derived based on such imputed components.

For categorical endpoints (such as IGA response), the above imputation will be performed for the respective scores on a continuous scale, then rounded to the nearest integer prior to deriving the response or resolution.

For the composite binary endpoints (such as ACR 20), the above imputation will be performed on each component with missing data and then response status will be determined based on such imputed components.

The treatment comparisons between a guselkumab group versus the placebo group will be performed using the analysis method specified for each of the N imputation datasets. The analysis results from all the N imputation datasets will be combined according to Rubin (Rubin 1987) and the p-value for testing the treatment difference will be obtained.

5.3.2.2. Imputing Missing Data Evaluating Deviation from Assumptions (Tipping Point Analyses)

5.3.2.2.1. For Primary Endpoint

For the **Primary endpoint** (ACR 20 response at Week 24) using the Adjusted Composite Estimand (see Section 5.3.4.2 for definition of estimand), the exhaustive scenario tipping point analyses will be performed to evaluate any deviation from the imputation of: missing NOT due to Natural Disaster or Major Disruption (see Section 5.3.4.2) being imputed as non-responder, and missing due to Natural Disaster or Major Disruption being imputed using FCS MI under MAR assumptions.

Note that as part of the ICE strategies for the Adjusted Composite Estimand (see Section 5.3.4.2), participants who meet ICEs 1-3 prior to Week 24 are set to non-responder regardless of observed data, and participants who meet ICEs 4 or 5 prior to Week 24 are imputed using FCS MI regardless of observed data. These participants, combined with participants with observed data at Week 24 who did not meet any ICEs prior that timepoint, comprise the not-to-be-varied group for the Adjusted Composite Estimand, and their response rate remains static through the tipping point analysis.

Let T_A be the total number of imputed values to-be-varied in the Active arm, where i of them will be set to 'Yes' response and (T_A-i) of them set to 'No' response. In the same vein, let T_P be the total number of imputed values to-be-varied in the Placebo arm, where j of them will be set to 'Yes' response and (T_P-j) of them set to 'No' response. The range of i is from 0 to T_A, and a range of j is from 0 to T_P, which is an 'exhaustive approach'.

Also for the **Primary endpoints** but using the Treatment Policy Estimand (see Section 5.3.4.2 for definition of estimand), the tipping point analyses based on imputed data by FCS MI will be performed to evaluate the impact of imputed when deviating from the MAR assumption.

- A pair of deltas (eg, Dg =-0.1, Dp =0.2) will be added to the predicted response rates of each missing data from the MI method depending on guselkumab or placebo group.
- With the new response rate, the missing response will be imputed for N (eg, N=200) times to generate N multiple imputations based on a Bernoulli distribution. Treatment comparisons will then be performed same as treatment comparison with MI.
- The range of delta values includes the scenarios where participants on guselkumab have worse outcomes than participants on placebo.

The Treatment Policy Estimand considers the ICEs irrelevant to the endpoint, thus all the missing data in the specified analysis population will be imputed and will comprise the to-be-varied group.

5.3.2.2.2. For Major Secondary Endpoints

For the **major secondary binary endpoints**, using the Treatment Policy Estimand (see Section 5.3.4.2 for definition of estimand), the tipping point analyses based on imputed data by FCS MI will be performed to evaluate the impact of imputed when deviating from the MAR assumption.

Note that as part of the ICE strategies for the Treatment Policy Estimand (see Section 5.3.4.2), the observed response for all participants will be used regardless of whether or not ICE criteria are met prior to Week 24. These participants comprise the not-to-be-varied group, and their response rates remain static through the tipping point analysis. The method of imputation is same with the tipping point analyses for Primary endpoints using the Treatment Policy Estimand.

For the **Major Secondary continuous endpoints** (change from baseline to Week 24 in HAQ-DI/SF-36 PCS/FACIT-F score) using the Treatment Policy Estimand (see Section 5.3.4.2), the tipping point analyses based on imputed data by FCS MI will be performed to evaluate the impact of imputed data when deviating from the MAR assumption.

- A delta (eg, Dg=0.2, Dp=0.1) will be added to the imputed value for each participant with missing value from the MI depending on whether the participant is in the guselkumab or placebo group.
- With the new datasets, treatment comparisons will be performed similar to treatment comparisons with MI data.
- The analysis will be repeated for a range of Dg and Dp by varying Dg and Dp independently, including the scenarios where participants on guselkumab have worse outcomes than participants on placebo.

5.3.3. Analysis Specifications

5.3.3.1. Level of Significance

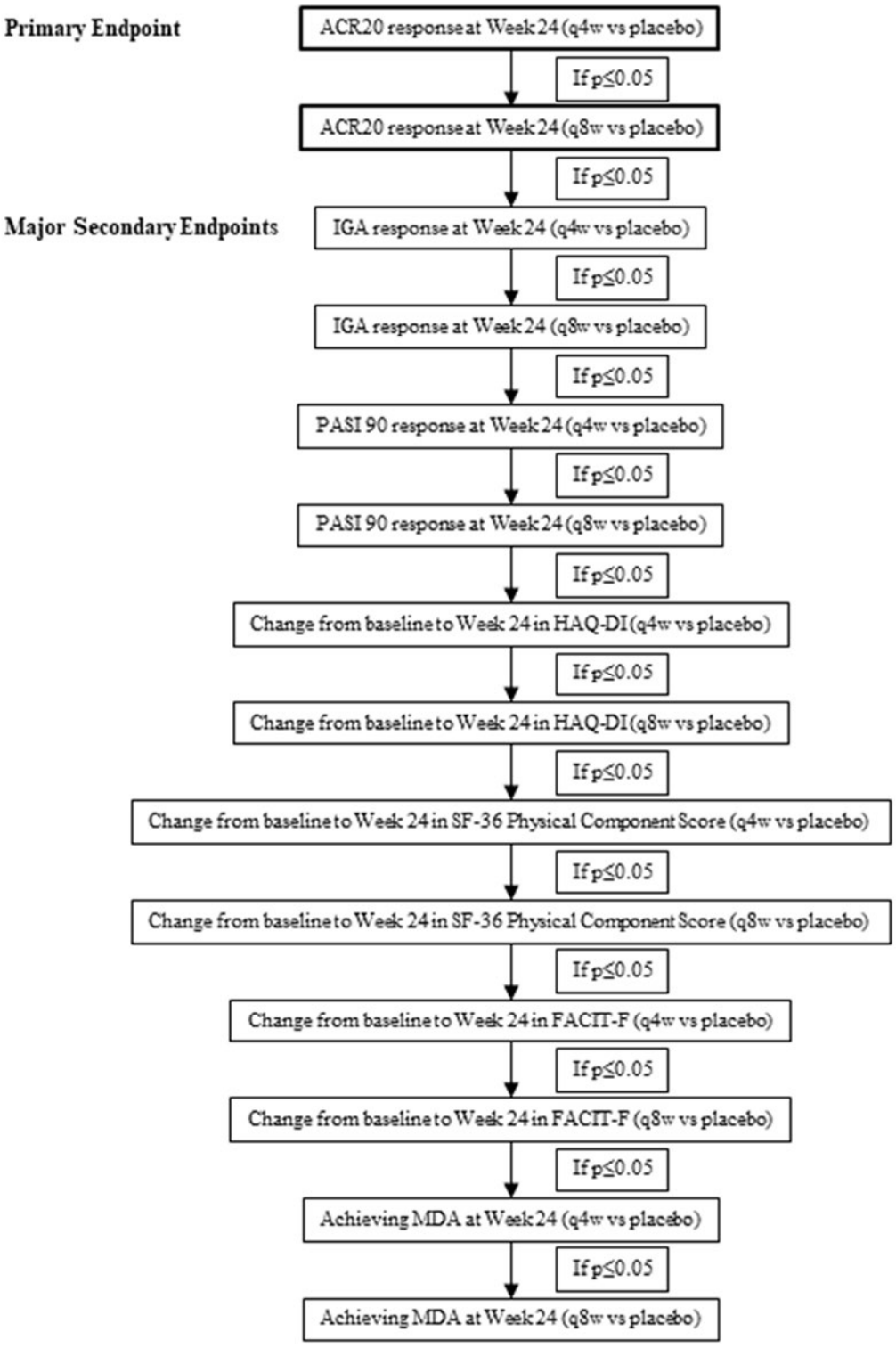
In general, statistical testing will be performed using 2-sided tests. The overall Type I error will be controlled among the primary and major secondary endpoints at 5%.

5.3.3.2. Multiplicity Adjustment for Testing Procedures

This study has 1 primary endpoint (proportion of participants who achieved an ACR 20 response at Week 24) and 10 major secondary endpoints. With 11 endpoints and 2 treatment comparisons for each of these endpoints, there are a total of 22 hypotheses to be tested. These hypotheses are explicitly listed in Section 6.11.

The overall Type I error of the 22 hypotheses will be controlled at a significant level of ≤0.05. The testing procedure tests the primary and selected major secondary endpoints, for the two regimens of guselkumab vs placebo, in a fixed sequence and each endpoint is tested at the two-sided 0.05 level of significance. The fixed sequence testing method tests an endpoint only if the null hypotheses of no difference between the gulsekumab regimen and placebo were rejected at the 0.05 level for all the endpoints above it in the sequence. This is shown visually in Figure 2 below.

Figure 2: Multiplicity Control



For the weakly controlled hypothesis tests (see Section 3.2), p-values are not strictly controlled. However, as its test depends on the positive result of its correlated strongly controlled hypothesis test, there are some levels of control even though the level of control is not quantifiable. Although

the level of control is not quantifiable, these reported p-values are not considered nominal. As a result, p-value for the weakly controlled hypothesis test will be reported with a footnote. The footnote will state that the test is not part of the sequential testing procedures but was prespecified to be tested upon achieving statistical significance of the strongly controlled primary endpoints.

5.3.4. Primary Endpoint Analysis

The primary endpoint of this study is the proportion of participants who achieved an ACR 20 response at Week 24. This section outlines the definitions and analyses of this primary endpoint.

5.3.4.1. Definition of Endpoint

ACR response is a composite measurement of change in PsA signs and symptoms and is presented as the numerical measurement of improvement in multiple disease assessment criteria (Felson 1993, 1995). An ACR20 response is defined as:

1. ≥20% improvement from baseline in both tender joint count (68 joints) [TJC68] and swollen joint count (66 joints) [SJC66]

   AND

2. ≥20% improvement from baseline in at least 3 of the following 5 assessments:
   a. Patient's Assessment of Pain (Visual analog scale [VAS]) [PAIN]
   b. Patient's Global Assessment of Disease Activity (arthritis, VAS) [GDPT]
   c. Physician's Global Assessment of Disease Activity (VAS) [GDEV]
   d. Patient's Assessment of Physical Function as measured by HAQ-DI
   e. C-reactive protein (CRP)

Following are the definitions of each of the forgoing disease assessment criteria (components) that are used in the determination of ACR20 response:

1. Tender Joint Count 68 (TJC68): a total number of tender joints among the 68 joints evaluated for tenderness. Each of the 68 joints will be evaluated for tenderness, categorized as tender or not tender. Joint evaluability rules specified in Section 6.10 for overwriting joint evaluation will be applied to those joints with joint injection(s)/surgical joint procedure(s). For participants with any joint not evaluable in the 68 joint set, joint count adjustment rules described in Section 6.10 will be applied in determining the ultimate count of tender joints.
2. Swollen Joint Count 66 (SJC66): a total number of swollen joints among the 66 joints evaluated for swelling. (Note: The 2 hip joints are excluded from swelling assessment.) Each of the 66 joints will be evaluated for swelling, categorized as swollen or not swollen. Joint evaluability rules specified in Section 6.10 for overwriting joint evaluation will be applied to those joints with joint injection(s)/surgical joint procedure(s). For participants with any joint not evaluable in the 66 joint set, joint count adjustment rules described in Section 6.10 will be applied in determining the ultimate count of swollen joints.
3. Patient's Assessment of Pain (PAIN): a measure from 0 (no pain) to 10 (the worst possible pain) on a 10-unit VAS.

4. Patient's Global Assessment of Disease Activity (arthritis, GDPT): a measure from 0 (very well) to 10 (very poor) on a 10-unit VAS.
5. Physician's Global Assessment of Disease Activity (GDEV): a measure from 0 (no arthritis activity) to 10 (extremely active arthritis) on a 10-unit VAS.
6. HAQ-DI: a measure of difficulty a participant may have in accomplishing tasks in 8 functional areas. For additional details, please refer to the definition of HAQ-DI in Section 5.3.6.2.2.
7. C-reactive protein (CRP): a lab parameter measured in mg/dL. Lower limit of quantification (LLOQ) rule specified in Section 6.10 will be applied to values <LLOQ.

If a participant's baseline value for a component is zero (ie, no disease activity as measured by that component), the participant should be considered as not achieving 20% improvement from baseline for that component since there is no room for improvement.

5.3.4.2. Estimand

5.3.4.2.1. Adjusted Composite Estimand (*Primary*)

The ***primary*** analysis for the primary endpoint will be based on the Adjusted Composite Estimand. This estimand is defined by the 5 components:

- **Population:** Participants with active PsA who have inadequate response and/or intolerance to 1 prior anti-TNF
- **Treatment:**
  - Placebo
  - Guselkumab
- **Variable:**

  ACR20 composite binary response variable at Week 24, where a responder is defined as a participant who achieves ACR20 response at Week 24 and does not experience ICE categories 1 to 3 prior to that time, in the hypothetical situation where Natural Disaster or Major Disruption and associated ICE categories 4 and 5 did not occur.

- **Intercurrent Events (ICEs):**
  1. Discontinued study intervention injections due to any reason **except** due to Natural Disaster or Major Disruption.
  2. Initiated or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroids from baseline for PsA.
  3. Initiated protocol prohibited medications/ therapies for PsA.
  4. Discontinued study intervention injections due to Natural Disaster or Major Disruption.
  5. Severe treatment non-compliance due to Natural Disaster or Major Disruption defined as ≥2 doses of study intervention missed due to Natural Disaster or Major Disruption.
- **Population level summary**: difference in proportion of responders (as per definition of **Variable** above) between guselkumab group and placebo group.

*Note that in the context of ICEs and missing data, the following are defined:

1. **Natural Disaster**: site closure, site access restrictions, or lockdowns caused by COVID-19.
2. **Major Disruption**: the disruption involving Ukraine and neighboring countries/territories beginning 24 February 2022.

This estimand examines the difference in proportion of participants who achieve ACR20 response after Week 24 without increasing/initiating select background PsA medications or discontinuing study intervention (for reasons not due to Natural Disaster or Major Disruption) prior to that point and in the hypothetical scenario where the Natural Disaster or Major Disruptions did not occur, between each guselkumab group individually versus placebo, amongst participants who had an IR and/or intolerance to one prior anti-TNF.

Additional supplementary analyses will be conducted on the primary endpoint using Treatment Policy Estimand, aiming to achieve a robust treatment effect for regulatory decision making.

5.3.4.2.2. Treatment Policy Estimand (*Supplementary*)

The treatment policy strategy is to use all observed data collected for the endpoint. The occurrence of the intercurrent event (ICE) is irrelevant: the value for the variable of interest is used regardless of whether or not the ICE occurs. This is the supplementary strategy aiming to achieve a robust treatment effect for regulatory decision making for primary and some major secondary endpoints.

This estimand is defined by the 5 components:

- **Population**: same as Adjusted Composite Estimand
- **Treatment**: same as Adjusted Composite Estimand
- **Variable**: ACR20 composite binary response variable at Week 24, where a responder is defined as a participant who achieves ACR20 response at Week 24 irrespective of background PsA medication or adherence to study intervention
- **Intercurrent Events**: the definitions of the 5 categories of ICE are same as the Adjusted Composite Estimand
- **Population level summary**: difference in proportion of responders (as per definition of **Variable** above) between guselkumab group and placebo group

This estimand examines the difference in proportion of participants who achieve ACR20 response after 24 weeks irrespective of background PsA medication or adherence to study intervention, between each guselkumab group individually versus placebo, amongst participants with active PsA who had an IR and/or intolerance to one prior anti-TNF.

5.3.4.3. Analysis Methods

The primary efficacy analysis of the primary endpoint will be analyzed at Week-24 DBL based on the **Adjusted Composite Estimand**.

In the primary efficacy analysis, data from all participants in FAS (Section 4.2) will be analyzed according to randomized intervention group regardless of the treatment actually received.

ICE Strategies

For the Adjusted Composite Estimand, ICEs 1-3 are incorporated as part of the endpoint using the composite strategy. This estimand acknowledges that a participant increasing their background PsA medications or electing to discontinue study intervention for reasons other than Natural Disaster or Major Disruption prior to the assessment timepoint is an unfavorable outcome, and thus for the purpose of analysis they are considered treatment failures (TF) and set to ACR20 non-responders at Week 24.

This estimand also possesses a hypothetical component, regarding significant unplanned changes to study conduct as a result of Natural Disaster or Major Disruption. This estimand seeks to estimate the treatment effect of study intervention as if the above, as well as ICEs directly resulting from it, did not occur. Thus, data observed after the occurrence of ICEs 4 and 5 will not be used and will be considered MAR and imputed using FCS MI on the individual ACR components.

Should there be a participant who meets ICEs from both the composite strategy and the hypothetical strategy, the composite strategy has precedence, and they will be set to an ACR20 non-responder at Week 24.

Handling Rules for Missing Data

After ICE strategies have been implemented, remaining missing data will be handled as follows:

1. **Missing Week 24 ACR 20 response due to Natural Disaster or Major Disruption,** will be assumed to be MAR and imputed using FCS MI on the individual ACR components.
2. **Missing Week 24 ACR 20 response NOT due to Natural Disaster or Major Disruption,** will be considered MAR but imputed conservatively as an ACR 20 non-responder (NRI rule) at Week 24 to be consistent with historical studies.

Analysis Testing

The treatment difference between each guselkumab group versus the placebo group will be tested using a CMH test stratified by baseline use of non-biologic DMARDs (yes, no) for each imputation set, and the Wilson-Hilferty (Ratitch 2013) transformation will be applied to the CMH statistics across the imputation sets. The transformed CMH statistics will be combined to calculate the p-values according to Rubin (Rubin 1987). The magnitude of the treatment difference will be estimated by the difference in ACR 20 response rates between the guselkumab and placebo groups with a 95% CI calculated based on Wald statistics (Kim 2013).

In order to control the overall Type 1 error rate, the primary endpoint will be tested in a fixed sequence.

1. Guselkumab 100 mg q4w versus placebo in ACR 20 response at Week 24, among FAS participants
2. Guselkumab 100 mg at Weeks 0, 4, and then q8w versus placebo in ACR 20 response at Week 24, among FAS participants

If the first test is significant at a 2-sided α-level of 0.05, the study will be considered positive, and the second test can then be performed.

5.3.4.4. Sensitivity and Supplementary Analyses

1. (**Sensitivity Analysis 1**) To evaluate the robustness of the **Adjusted Composite Estimand** regarding the assumptions for missing or not used data, sensitivity analyses with the exhaustive two-dimensional scenario tipping point analyses will be performed. The analysis will be conducted for an 'exhaustive approach' testing all combinations of missing data imputation as responder and non-responder (NR) (Section 5.3.2.2). Note that data imputed as part of the ICE strategy (ie, NR due to meeting ICEs 1-3, FCS MI due to meeting ICEs 4 and 5) will be performed prior to the tipping point analysis. The chi-square test will be used to compare each guselkumab group versus the placebo group. This will avoid the complication of having to incorporate baseline stratification in the mix when generating all combinations of responders and NR for the missing data for CMH test. As all combinations will be presented, both the points where tipping occur, as well as the proportion of non-tipping combinations, are of interest.

2. (**Sensitivity Analysis2**) To assess the effect of excluding all Ukrainian participants from the analysis for the **Adjusted Composite Estimand**, as all sites in Ukraine have been impacted by Major Disruption (eg, study intervention interruptions, temporary site closures, etc) to a larger or smaller extent. The same analysis of ACR 20 response at Week 24 as the primary analysis will be performed on the FAS-UKR analysis set.

3. (**Sensitivity Analysis 3**) To assess the effect of protocol deviations which may affect efficacy, a sensitivity analysis similar to the main analysis will be performed but based on the PPAS.

4. (**Sensitivity Analysis** 4) As the Major Disruption has the potential to prevent source data verification (SDV) from taking place (eg, due to site closure, site inaccessibility, etc), an analysis will be conducted to assess the impact by excluding affected participants. The same analysis as the main analysis will be performed on the FAS analysis set, excluding participants for which the expected SDV was not completed for data on or prior to Week 24.

5. (**Supplementary Analysis 1**) To support regulatory decision making, the **Treatment Policy Estimand** (Section 5.3.4.2) will also be evaluated as a supplementary analysis. In this analysis, the observed ACR 20 response for all participants will be used regardless of whether or not ICE criteria are met prior to Week 24, and the missing ACR 20 response will be imputed by FCS MI (Section 5.3.2.1) on the component level under the assumption that data are MAR. Treatment comparisons for each imputation data set will be based on a CMH test stratified by baseline use of non-biologic DMARDs (yes, no), as per Section 5.3.1. This analysis will be conducted based on the FAS-UKR analysis set.

6. (**Supplementary Analysis 2**) An additional tipping point analysis for the **Treatment Policy Estimand** will be performed. The observed ACR 20 response for participants will be used regardless of whether or not ICE criteria are met prior to Week 24, and the missing ACR 20 response will be imputed by FCS MI (Section 5.3.2.1) on the component level under the assumption that data are MAR. In this analysis, a pair of deltas will be added to the predicted response rates from MI method depending on guselkumab or placebo group to new MI datasets (Section 5.3.2) to evaluate deviation from the assumption of MAR for missing data. The same analysis method as in the primary analysis will be applied for the pairs of deltas. The analysis will be done for pairs of delta values include the scenarios where participants on

guselkumab have worse outcomes than participants on placebo. This analysis will be conducted based on the FAS-UKR analysis set.

Subgroup Analyses

Subgroup analyses will be performed using the Adjusted Composite Estimand using a logistic regression model on the multiply imputed data to evaluate treatment consistency in proportion of participants who achieve an ACR 20 response at Week 24 over baseline demographics, baseline disease characteristics, and prior and baseline medication use. A forest plot will be produced for all subgroups listed in Section 5.5.5. The odds ratios and the corresponding 95% CIs will also be provided for each of the subgroups (Section 5.3.1). In addition, the p-values for interaction of the intervention groups and the subgroups will also be provided when a subgroup has at least 2 categories.

If the number of participants in a subgroup is too small (eg, <10), subgroups may be pooled for analyses.

5.3.4.5. Summary of Analyses Related to the Primary Endpoint of ACR 20 Response at Week 24

Table 6 below provides an overview on all the analyses related to the primary endpoint of ACR 20 response at Week 24, the estimands, the data handling rules to be used, and the analysis methods and summary statistics. Section 5.3.2 provides a summary of the MI method.

Table 6: Summary of Analyses Related to the Primary Endpoint of ACR 20 Response at Week 24

Analysis (Estimand) – Analysis Set	**TF/Missing data imputation**	**Additional notes**
Primary Analysis (based on Adjusted Composite Estimand) – FAS	ICEs 1-3 considered as TF ICEs 4 and 5, data not used, but rather imputed using FCS MI Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Summarized descriptively • Pooled response rates, treatment difference in response rates and 95% CI across multiply imputed data sets • P-value from CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets
Sensitivity Analysis 1 (based on Adjusted Composite Estimand) – FAS	ICEs 1-3 considered as TF ICEs 4 and 5, data not used, but rather imputed using FCS MI, then averaged across MI datasets. Missing data rules not applied. Response rate of missing data varied with all possible combinations of response status for missing data. Single imputation	• Exhaustive 2-dimensional tipping point analysis; graphical • The chi-squared test is used to compare intervention groups, for each coordinate on the graph separately

Analysis (Estimand) – Analysis Set	TF/Missing data imputation	Additional notes
Sensitivity Analysis 2 (based on Adjusted Composite Estimand) – FAS-UKR	ICEs 1-3 considered as TF ICEs 4 and 5, data not used, but rather imputed using FCS MI Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Summarized descriptively • Pooled response rates, treatment difference in response rates and 95% CI across multiply imputed data sets • P-value from CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets
Sensitivity Analysis 3 (based on Adjusted Composite Estimand) – PPAS	ICEs 1-3 considered as TF ICEs 4 and 5, data not used, but rather imputed using FCS MI Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Summarized descriptively • Pooled response rates, treatment difference in response rates and 95% CI across multiply imputed data sets • P-value from CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets
Sensitivity Analysis 4 (based on Adjusted Composite Estimand) – FAS, Excluding Participants with Non-SDV'd Data Through Week 24	ICEs 1-3 considered as TF ICEs 4 and 5, data not used, but rather imputed using FCS MI Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Summarized descriptively • Pooled response rates, treatment difference in response rates and 95% CI across multiply imputed data sets • P-value from CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets
Supplementary Analysis 1 (based on Treatment Policy Estimand) – FAS-UKR	No ICEs considered as TF Missing data assumed MAR, imputed via FCS MI	• Summarized descriptively • Pooled response rates, treatment difference in response rates and 95% CI across multiply imputed data sets • P-value from CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets

Analysis (Estimand) – Analysis Set	TF/Missing data imputation	Additional notes
Supplementary Analysis 2 (based on Treatment Policy Estimand) – FAS-UKR	No ICEs considered as TF Missing data imputed via FCS MI, then imputed data response rate varied and generated using Bernoulli distribution	• 2-dimensional tipping point analysis to assess robustness of analysis results should there be deviation from MAR assumption of missing data; graphical • P-value from CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets, for each coordinate on the graph separately
Subgroup Analyses (based on Adjusted Composite Estimand) – FAS subgroups	ICEs 1-3 considered as TF ICEs 4 and 5, data not used, but rather imputed using FCS MI Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Odds ratio and 95% CI for treatment comparison • P-value from logistic regression for the interaction of intervention group and subgroup variable • Graphical: forest plots

5.3.5. Major Secondary Analysis

The 10 major secondary analyses are:

1. Proportion of participants who achieve a psoriasis response of IGA psoriasis score of 0 (cleared) or 1 (minimal) AND ≥2-grade reduction from baseline at Week 24 among the participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline
2. Proportion of participants who achieve PASI 90 response at Week 24 among the participants with ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline
3. Change from baseline in HAQ-DI at Week 24
4. Change from baseline in the SF-36 PCS at Week 24
5. Change from baseline in FACIT-F score at Week 24
6. Proportion of participants achieving MDA at Week 24
7. Proportion of participants who achieve ACR20 response at Week 16
8. Proportion of participants who achieve ACR50 response at Week 16
9. Proportion of participants who achieve ACR50 response at Week 24
10. Proportion of participants who achieve ACR70 response at Week 24

This section outlines the definition and analyses of these major secondary endpoints. All the secondary endpoints will be analyzed at Week 24 DBL according to the randomized treatment groups. Data from all subjects in FAS (Section 4.2) will be included with the following exceptions: the analysis of the psoriasis response of IGA and PASI 90 will be based on FAS among the subjects with a ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline.

5.3.5.1. Estimand

5.3.5.1.1. Adjusted Composite Estimand ***(Primary)***

The ***primary*** analysis for the major secondary endpoints will be based on the Adjusted Composite Estimand. This estimand was previously defined for the primary endpoint at Week 24 in Section 5.3.4.2.1. The variables are generalized across endpoints below:

Variable:

- Binary: The endpoint (eg, ACR 20) is defined as responders who achieve response at Week 24 and do not experience ICE categories 1 to 3 prior to that time, in the hypothetical situation where Natural Disaster or Major Disruption and associated ICE categories 4 and 5 did not occur.
- Continuous: The endpoint is defined as change from baseline score prior to experiencing ICE categories 1 to 3 and 0 (no improvement) after experiencing ICE categories 1 to 3 regardless of observed data, and in the hypothetical situation where Natural Disaster or Major Disruption and associated ICE categories 4 and 5 did not occur.

5.3.5.1.2. Treatment Policy Estimand *(Supplementary)*

For the Treatment Policy Estimand, the occurrence of ICEs is considered irrelevant. This estimand looks at the effect of assignments on intervention groups irrespective of changes to background PsA medications, study intervention adherence, or study retention.

This estimand was previously defined for the primary endpoint at Week 24 in Section 5.3.4.2.2.

The variables are generalized across endpoints below:

Variable:

- Binary: The endpoint (eg, ACR 20) is defined as responders who achieve response at Week 24 irrespective of background PsA medication or adherence to study intervention
- Continuous: The endpoint is defined as change from baseline score irrespective of background PsA medication or adherence to study intervention

The Treatment Policy Estimand will be analyzed for all major secondary endpoints.

5.3.5.2. Proportion of Subjects Who Achieve a Psoriasis IGA Response at Week 24 Among the Subjects with ≥3% BSA Psoriatic Involvement and an IGA Score of ≥2 (mild) at Baseline

5.3.5.2.1. Definition

CCI [redacted]

CCI

[redacted]

.

5.3.5.2.2. Analysis Methods

Data from all subjects in FAS who had ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline will be included and analyzed according to the randomized treatment groups.

The same analysis method as described in Section 5.3.4.3 will be applied to the treatment comparisons on proportion of subjects with a psoriasis IGA response at Week 24.

To support regulatory decision making, a supplementary on proportion of subjects with a psoriasis IGA response analysis will be provided in FAS based on the Treatment Policy Estimand. Analysis of at Week 24 will be performed using an CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets similar to the main analysis.

5.3.5.3. Proportion of Participants Who Achieve PASI 90 Response at Week 24 Among the Participants With ≥3% BSA Psoriatic Involvement and an IGA Score of ≥2 (mild) at Baseline

5.3.5.3.1. Definition

CCI

[redacted]

CCI [redacted]

5.3.5.3.2. Analysis Methods

Data from all subjects in FAS who had ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline will be included and analyzed according to the randomized treatment groups.

The same analysis method as described in Section 5.3.5.2.2 will be applied to the treatment comparisons on proportion of subjects who achieve PASI 90 response at Week 24.

5.3.5.4. Change From Baseline in Health Assessment Questionnaire Disability Index (HAQ-DI) at Week 24

5.3.5.4.1. Definition

CCI [redacted]

5.3.5.4.2. Analysis Methods

The change from baseline in HAQ-DI score at Week 24 will be analyzed at Week 24 DBL based on the Adjusted Composite Estimand (Section 5.3.4.2). In this analysis, data from all subjects in

FAS (Section 4.2) will be analyzed according to randomized treatment group regardless of the treatment received.

Analysis of the change from baseline in HAQ-DI score at Week 24 will be performed using an ANCOVA model based on MI data. The MI method will be applied to impute for missing data, under the assumption of MAR, the missing change score of HAQ-DI from baseline at Week 24 defined in the Adjusted Composite Estimand (ie, change from baseline with no change assumed at and after meeting ICE (1-3) criteria). The treatment difference between each guselkumab group versus the placebo group will be tested for each imputation dataset and then the analysis results across all imputation datasets will be combined.

To support regulatory decision making, a supplementary analysis will be provided in FAS-URK based on the Treatment Policy Estimand. Analysis of the change from baseline in HAQ-DI score at Week 24 will be performed using an ANCOVA model based on MI data similar to the main analysis.

5.3.5.5. Change from Baseline in SF-36 PCS Score at Week 24

5.3.5.5.1. Definition

[redacted]

5.3.5.5.2. Analysis Methods

Data from all subjects in FAS will be included and analyzed according to the randomized treatment groups.

The same analysis methods as described in Section 5.3.5.4.2 will be applied to treatment comparisons on the change from baseline in SF-36 PCS at Week 24.

5.3.5.6. FACIT-Fatigue Questionnaire

CCI [redacted]

5.3.5.7. Proportion of Subjects with Minimal Disease Activity (MDA) at Week 24

CCI [redacted]

[redacted]

5.3.5.8. Proportion of Subjects with ACR 50 Response at Week and 24, Proportion of Subjects with ACR 20 Response at Week 16, Proportion of Subjects with ACR 50 Response at Week 16, and Proportion of Subjects with ACR 70 Response at Week 24

5.3.5.8.1. Definition

[redacted]

5.3.5.8.2. Analysis Methods

Data from all subjects in FAS (Section 4.2) will be analyzed according to randomized treatment group regardless of the treatment actually received.

The same analysis method as described in Section 5.3.5.2.2 will be applied to the treatment comparisons on proportion of subjects with ACR 50 response at Week 24, proportion of subjects with ACR 20 response at Week 16, proportion of subjects with ACR 50 response at Week 16, and proportion of subjects with ACR 70 response at Week 24.

5.3.5.9. Method of Analysis

In general, descriptive statistics, such as mean, SD, median, IQ range, minimum, and maximum for continuous variables, and counts and percentages for discrete variables will be used to summarize most data.

All the analyses will be based on the FAS (Section 4.2). Table 7 provides a summary of analyses related to major secondary endpoints.

The **main analysis** for Major Secondary Efficacy Endpoints will be conducted using the Adjusted Composite Estimand. The ICE Strategy handling rules for the estimand will be applied first, after which the data handling rules for missing data will be applied. Specifically:

ICE Strategies for Adjusted Composite Estimand

For the Adjusted Composite Estimand, ICE categories 1-3 are incorporated as part of the endpoint using the composite strategy. This estimand acknowledges that a participant increasing their

background PsA medications, electing to discontinue study intervention for reasons other than Natural Disaster or Major Disruption is an unfavorable outcome where participants who meet them prior to the visit will be considered as treatment failures at that visit and subsequently through the final efficacy visit. Participants meeting ICE categories 1-3 are considered non-responders for binary response endpoints, and to have no improvement (change from baseline=0) for continuous endpoints, regardless of observed data.

This estimand also employs the hypothetical strategy. Primarily, regarding significant unplanned changes to study conduct as a result of Natural Disaster or Major Disruption. This estimand seeks to estimate the treatment effect of study intervention as if Natural Disaster or Major Disruption, as well as ICEs directly resulting from them, did not occur. Thus, data observed after the occurrence of ICEs 4 and 5 will not be used and will be considered MAR and imputed using FCS MI.

For participants experiencing multiple ICEs, an ICE in categories 1-3 (ie, using the composite strategy) will supersede an ICE in categories 4 or 5 (ie, using the hypothetical strategy).

Handling Rules for Missing Continuous Data for Adjusted Composite Estimand

1. **Missing data for any reason** will be assumed to be MAR and imputed using FCS MI on continuous component scale.

Handling Rules for Missing Binary Data for Adjusted Composite Estimand

1. Missing data due to Natural Disaster (caused site closure, site access restrictions, or lockdowns) or Major Disruption will be assumed to be MAR and imputed using FCS MI on the individual components.
2. Missing data for any other reason will be considered MAR but imputed conservatively as a non-responder (NRI rule) at Week 24 to be consistent with historic.

Analysis Testing for Adjusted Composite Estimand

Statistical comparison between a guselkumab group (100 mg q4w or 100 mg q8w) and the placebo group will be performed at Week 24 using ANCOVA model treatment difference across multiply imputed datasets for continuous endpoints, and the CMH with Wilson-Hilferty transformation across multiply imputed data sets for treatment difference and 95% CI will be used for the binary response endpoints.

5.3.5.10. Additional Tipping Point Analyses

To support regulatory decision making, additional tipping point analyses will be performed for each major secondary endpoint to evaluate the impact when the missing data deviate from the MAR assumption. Analyses will be performed for the Treatment Policy Estimand (Section 5.3.4.2), where observed data is included regardless of whether or not ICE criteria are met prior to Week 24 and the missing will be imputed using FCS MI.

ICE Strategies for Treatment Policy Estimand

For the Treatment Policy Estimand, the occurrence of ICEs is considered irrelevant. This estimand looks at the effect of assignment to intervention group irrespective of changes to background PsA medications, study intervention adherence, or study retention.

Handling Rules for Missing Data

1. Missing binary data for any reasons will first be imputed using MI, then a pair of deltas will be added to the predicted response or resolution rates of each missing subject from the MI method depending on guselkumab or placebo group to new MI datasets (Section 5.3.2.2).
2. Missing continuous data for any reasons will first be imputed using MI, then a pair of deltas will be added to the imputed values from the MI method depending on guselkumab or placebo group (Section 5.3.2.2).

Analysis Testing for Treatment Policy Estimand

The same analysis method as in supplementary analysis 2 will be applied for the pairs of deltas. The treatment comparison will use the CMH test for binary variables and ANCOVA model for continuous variables. The analysis results will be presented graphically.

Table 7 provides an overview of all the analyses related to the major secondary endpoints, the estimands, the data handling rules to be used, and the analysis methods and summary statistics.

Table 7: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Major Secondary Endpoints

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
ENDPOINTS AT WEEK-24 DBL, ADJUSTED COMPOSITE ESTIMAND				
1	• Proportions of participants with ACR 20 response at Week16 • Proportions of participants with ACR 50 response at Week16/Week24 • Proportions of participants with ACR 70 response at Week24	FAS	Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Summarized descriptively • Response rates, treatment difference in response rates and 95% CI • p-value from the CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets
2	Change from baseline in HAQ-DI score / SF-36 PCS score / FACIT-F score at Week24	FAS	Missing data for any reason assumed MAR, imputed via FCS MI	• Summarized descriptively • ANCOVA for each MI dataset and then analysis results across all MI datasets combined • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) between groups, p-values for differences between groups

3	Proportion of participants with IGA response / PASI90 response	FAS with ≥3% BSA and IGA≥2 at baseline	Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Summarized descriptively • Response rates, treatment difference in response rates and 95% CI • p-value from the CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets
4	Proportion of participants with MDA response	FAS	Missing data due to Natural Disaster or Major Disruption assumed MAR, imputed via FCS MI Missing data not due to Natural Disaster or Major Disruption assumed MAR but conservatively imputed as non-response	• Summarized descriptively • Response rates, treatment difference in response rates and 95% CI • p-value from the CMH statistic with Wilson-Hilferty transformation across multiply imputed data sets

ENDPOINTS AT WEEK-24 DBL USING TIPPING POINT ANALYSIS WITH TREATMENT POLICY ESTIMAND AND The FAS-UKR ANALYSIS SET INSTEAD OF FAS

The endpoints analyzed using the Adjusted Composite Estimand for corresponding visits/DBL are repeated here, except:

- All observed data will be used, regardless of whether or not the subjects meet any ICEs.
- Missing data assumed MAR, imputed via FCS MI regression.
- Two-dimensional tipping point analysis will be used.
- Continuous endpoints will use ANCOVA and binary response endpoints will use CMH statistic to compare treatment groups.
- Analysis results to be presented graphically.

5.3.6. Other Efficacy Endpoints

In addition to the primary and secondary major endpoints, CCI [redacted]

CCI [redacted]

5.3.6.1. Estimands

The estimand is composed of 5 components: Population, treatment, variable, ICEs, and population level summary. Of these, population, variable, and population level summary vary across the different endpoints, while treatment and ICEs are constant. For all endpoints, the treatments are placebo and the 2 guselkumab dose groups. Additionally, the definition of ICEs is defined below.

Adjusted Composite Estimand *(Primary)*

The ***primary*** analysis for the Other Efficacy endpoints will be based on the Adjusted Composite Estimand. This estimand was previously defined for the primary analysis of ACR 20 response at Week 24 in Section 5.3.4.2 and for the change from baseline score at Week 24 in Section 5.3.5.1. The ICEs are generalized across endpoints and for all efficacy visits below:

- **Intercurrent Events:**
  1. Discontinued study intervention injections due to any reason **except** due to Natural Disaster or Major Disruption.
  2. Initiated or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroids over baseline for PsA.
  3. Initiated protocol prohibited medications/ therapies for PsA.
  4. Discontinued study intervention injections due to Natural Disaster or Major Disruption.
  5. Severe treatment non-compliance due to Natural Disaster or Major Disruption. This is defined for a given visit, when the total number of doses of study intervention missed due to Natural Disaster or Major Disruption exceeds 30% of the total protocol defined doses from Week 0 up to and including that visit. For Week 20, this amounts to ≥2 dose missed.

*Note that in the context of ICEs and missing data, the following are defined:

1. **Natural Disaster**: site closure, site access restrictions, or lockdowns caused by COVID-19.
2. **Major Disruption**: the disruption involving Ukraine and neighboring countries/territories beginning 24 February 2022.

Treatment Policy Estimand *(Supplementary)*

This estimand was previously defined for the primary endpoint of ACR 20 response at Week 24 in Section 5.3.4.2 and for the change from baseline score at Week 24 in Section 5.3.5.1. The ICEs and their handling are generalized across endpoints and for all efficacy visits below:

- **Intercurrent Events**: the definition of the 5 categories of ICE is the same as the generalized Adjusted Composite Estimand in this section.

5.3.6.2. Endpoints Related to Reduction of Signs and Symptoms and Physical Function

In this study, Other Endpoints for reduction of signs and symptoms and physical function include those related to CCI [redacted]

All Other Endpoints related to reduction of signs and symptoms and physical function will be conducted under th CCI [redacted]

5.3.6.2.1. CCI Related Endpoints

CCI

The CCI related Other Endpoints include:

CCI

5.3.6.2.2. CCI Related Endpoints

CCI

The CCI related Other Endpoints include:

CCI

.

CCI [redacted]

5.3.6.2.3. CCI [redacted] Related Endpoints

CCI [redacted]

CCI [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

The CCI [redacted] related Other Endpoints include:

CCI [redacted]

[redacted].

CCI [redacted]

CCI [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

5.3.6.2.4. CCI Related Endpoints

[redacted]

The CCI related Other Endpoints include:

[redacted]

5.3.6.2.5. CCI Related Endpoints

[redacted]

CCI

The CCI related to Other Endpoints include:

CC

5.3.6.2.6. CCI Related Endpoints

CCI

CCI [redacted]

The CCI [redacted] related Other Endpoints include:

C
C [redacted]

5.3.6.2.7. CCI [redacted] Related Endpoints

CCI [redacted]

CCI

The CCI related Other Endpoints include:

CCI

5.3.6.2.8. CCI CCI

CCI

The CCI and CCI related Other Endpoints include:

CCI

5.3.6.2.9. Method of Analysis

In general, descriptive statistics, such as mean, SD, median, IQ range, minimum, and maximum for continuous variables, and counts and percentages for discrete variables will be used to summarize most data.

The **main analysis** for Other Efficacy endpoints related to reduction of signs and symptoms and physical function will be conducted using the Adjusted Composite Estimand. The ICE Strategy handling rules for the estimand will be applied first, after which the data handling rules for missing data will be applied. Specifically:

ICE Strategies for Adjusted Composite Estimand

For the Adjusted Composite Estimand, ICE categories 1, 2, 3 are incorporated as part of the endpoint using the composite strategy. This estimand acknowledges that a participant increasing their background PsA medications, electing to discontinue study intervention for reasons other than Natural Disaster or Major Disruption is an unfavorable outcome where participants who meet them prior to the visit will be considered as treatment failures at that visit and subsequently through the final efficacy visit. Participants meeting TF are considered non-responders for binary response endpoints, and to have no improvement (change from baseline=0) for continuous endpoints, regardless of observed data.

This estimand also employs the hypothetical strategy. Primarily, regarding significant unplanned changes to study conduct as a result of Natural Disaster or Major Disruption. This estimand seeks to estimate the treatment effect of study intervention as if Natural Disaster or Major Disruption, as well as ICEs directly resulting from them, did not occur. Thus, for ICE category 4 all observed data after meeting the ICE through end of study will not be used and will be assumed to be MAR. For ICE category 5, observed data at the visit immediately subsequent to meeting the ICE will not be used and will be assumed to be MAR.

For participants experiencing multiple ICEs, an ICE in categories 1, 2, 3 (ie, using the composite strategy) will supersede an ICE in categories 4 or 5 (ie, using the hypothetical strategy).

These ICE strategies are also summarized in tabular form in Appendix 14 in Section 6.14.

Handling Rules for Missing Continuous Data for Adjusted Composite Estimand

1. **Missing data for any reason** will be assumed to be MAR. The data is not explicitly imputed but is accounted for in the analysis model.

Handling Rules for Missing Binary Data for Adjusted Composite Estimand

1. Missing data due to Natural Disaster (caused site closure, site access restrictions, or lockdowns) or Major Disruption will be assumed to be MAR. The data is not explicitly imputed but is accounted for in the analysis model.
2. Missing data for any other reason will be assumed to be MAR but conservatively imputed as NR.

Analysis Testing for Adjusted Composite Estimand

Statistical comparison between a guselkumab group (100 mg q4w or 100 mg q8w) and the placebo group will be performed by visit through Week 24 using the MMRM model for continuous endpoints and the GLMM for binary response endpoints (see Appendix 13 in Section 6.13). No treatment comparison will be performed after Week 24. For continuous endpoints, only descriptive summaries and LSmeans by study intervention and visit will be presented after Week 24 through Week 100; for binary endpoints, only descriptive summaries and model-based response rates by study intervention and visit will be presented after Week 24 through Week 100.

ICE Strategies for Treatment Policy Estimand

For the Treatment Policy Estimand, the occurrence of ICEs is considered irrelevant. This estimand looks at the effect of assignment to intervention group irrespective of changes to background PsA medications, study intervention adherence, or study retention.

Handling Rules for Missing Data for Treatment Policy Estimand

1. **Missing data for any reason** will not be imputed, assumed to be MCAR.

Analysis Testing for Treatment Policy Estimand

A corresponding supportive analysis will be conducted using the Treatment Policy Estimand. For these analyses, the FAS-UKR analysis set or further subset (if specified) will be used. For visits through Week 24, an ANCOVA model will be used for continuous endpoints, and the CMH test for treatment difference and 95% CI calculated based on Wald statistics will be used for the binary response endpoints. No treatment comparison will be performed after Week 24. For continuous endpoints, only descriptive summaries and LSmeans by study intervention and visit will be presented after Week 24 through Week 100; for binary endpoints, only descriptive summaries of response rates by study intervention and visit will be presented after Week 24 through Week 100.

Table 8 summarizes the analyses for supportive efficacy endpoints related to reduction of signs and symptoms and physical function, the methods for analyses, and the data handling rules used.

Table 8: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Endpoints of Signs & Symptoms and Physical Function

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
ENDPOINTS BY VISIT THROUGH WEEK 24 AT WEEK-24 DBL, ADJUSTED COMPOSITE ESTIMAND				
1	Proportions of participants with ACR 20, ACR 50, and ACR 70 response	FAS	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates, and treatment difference in response rates and 95% CI and p-value, based on GLMM
2	Change from baseline in ACR components	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group
3	Change from baseline in HAQ-DI score	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
4	Proportion of participants with HAQ-DI response	FAS whose baseline HAQ-DI score ≥0.35	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates, and treatment difference in response rates and 95% CI and p-value, based on GLMM

Table 8: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Endpoints of Signs & Symptoms and Physical Function

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
5	Proportion of participants who achieve resolution of enthesitis	FAS with enthesitis at baseline	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates, and treatment difference in response rates and 95% CI and p-value, based on GLMM
6	Proportion of participants with resolution of dactylitis	FAS with dactylitis at baseline	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates, and treatment difference in response rates and 95% CI and p-value, based on GLMM
7	Change from baseline in enthesitis score (based on Leeds Enthesitis Index [LEI])	FAS with enthesitis at baseline	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
8	Change from baseline in enthesitis score (based on Spondyloarthritis Research Consortium of Canada [SPARCC])	FAS with enthesitis at baseline	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
9	Change from baseline in dactylitis score	FAS with dactylitis at baseline	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
10	Proportion of participants with MDA	FAS	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates, and treatment difference in response rates and 95% CI and p-value, based on GLMM
11	Proportion of participants with VLDA	FAS	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates, and treatment difference in response rates and 95% CI and p-value, based on GLMM
12	Change from baseline in DAPSA score	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM

Table 8: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Endpoints of Signs & Symptoms and Physical Function

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
13	Change from baseline in PASDAS	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
14	Change from baseline in BASDAI	FAS with spondylitis and peripheral joint involvement as their primary arthritic presentation of PsA at baseline	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
'-' indicates no missing data rules to be applied.				
ENDPOINTS BY VISIT AFTER WEEK 24 THROUGH WEEK 52 AT WEEK-52 DBL, ADJUSTED COMPOSITE ESTIMAND				
1, 4, 5, 6, 10, 11	Same as through Week 24	Same as through Week 24	Same as through Week 24	• Summarized descriptively • Response rates by intervention group based on GLMM
2, 3, 7, 8, 9, 12, 13, 14	Same as through Week 24	Same as through Week 24	Same as through Week 24	• Summarized descriptively • LSMeans and 95% CI by intervention group based on MMRM model
15	Proportion of participants who maintained an ACR 20 response at Week 52	FAS who achieved an ACR 20 response at Week 24	NRI for MAR not due to Natural Disaster or Major Disruption.	• Summarized descriptively • Response rates by intervention group based on GLMM
16	Proportion of participants who maintained an ACR 50 response at Week 52	FAS who achieved an ACR 50 response at Week 24	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates by intervention group based on GLMM
17	Proportion of participants who maintained an ACR 70 response at Week 52	FAS who achieved an ACR 70 response at Week 24	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates by intervention group based on GLMM
18	Proportion of participants who maintained HAQ-DI response at Week 52	FAS who achieved HAQ-DI response at Week 24	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates by intervention group based on GLMM
ENDPOINTS BY VISIT AFTER WEEK 52 THROUGH WEEK 100 AT WEEK-112 DBL, ADJUSTED COMPOSITE ESTIMAND				

Table 8: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Endpoints of Signs & Symptoms and Physical Function

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
1, 4, 5, 6, 10, 11, 15, 16, 17, 18	Same as through Week 52	Same as through Week 52	Same as through Week 52	• Summarized descriptively • Response rates by intervention group based on GLMM
2, 3, 7, 8, 9, 12, 13, 14	Same as through Week 52	Same as through Week 52	Same as through Week 52	• Summarized descriptively • LSMeans and 95% CI by intervention group based on MMRM model
ENDPOINTS BY VISIT THROUGH WEEK 24 AT WEEK-24 DBL, AFTER WEEK 24 THROUGH WEEK 52 AT WEEK-52 DBL, AFTER WEEK 52 THROUGH WEEK 112 AT WEEK-112 DBL, USING TREATMENT POLICY ESTIMAND AND THE FAS-UKR ANALYSIS SET INSTEAD OF FAS				
The endpoints analyzed using the Adjusted Composite Estimand for corresponding visits/DBL are repeated here, except: • No imputation will be applied, and all analyses will be based on observed data. • Continuous endpoints will use ANCOVA instead of MMRM; binary response endpoints will use CMH test with CI based on Wald statistics instead of GLMM for visits up to and including Week 24.				

5.3.6.3. Endpoints Related to Skin Disease

In this study, Other Endpoints for skin disease include those related to CCI [redacted]

CCI [redacted]

5.3.6.3.1. CCI [redacted] Related Endpoints

CCI [redacted]

[redacted]

[redacted]

- CCI [redacted]

[redacted]

CCI [redacted]

[redacted]

The CCI [redacted] related Other Endpoints include:

CCI [redacted]

5.3.6.3.2. CCI [redacted] Related Endpoints

CCI [redacted]

[redacted]

The CCI [redacted] related Other Endpoints include:

[redacted]

5.3.6.3.3. Method of Analysis

In general, the same methods of analysis for the Other Endpoints related to reduction of signs and symptoms and physical function in Section 5.3.6.2.9 will also be used for those related to skin disease.

Table 9 summarizes the analyses for supportive efficacy endpoints related to skin disease, the methods for analyses, and the data handling rules used.

Table 9: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Endpoints of Skin Disease

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
ENDPOINTS BY VISIT THROUGH WEEK 24 AT WEEK-24 DBL, ADJUSTED COMPOSITE ESTIMAND				
1	Proportions of participants with IGA score of 0	FAS with ≥3% BSA and IGA≥2 at baseline	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates by intervention group based on GLMM
2	Proportions of participants with PASI ≥75%, ≥90%, ≥100% improvement	FAS with ≥3% BSA and IGA≥2 at baseline	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates by intervention group based on GLMM
3	Change from baseline in PASI score	FAS with ≥3% BSA and IGA≥2 at baseline	-	• Summarized descriptively • LSMeans and 95% CI by intervention group based on MMRM model
'-' indicates no missing data rules to be applied				
ENDPOINTS BY VISIT AFTER WEEK 52 THROUGH WEEK 100 AT WEEK-112 DBL, USING ADJUSTED COMPOSITE ESTIMAND				
1, 2	Same as through Week 24	Same as through Week 52	Same as through Week 52	• Summarized descriptively • Response rates by intervention group based on GLMM
3	Same as through Week 24	Same as through Week 52	Same as through Week 52	• Summarized descriptively • LSMeans and 95% CI by intervention group based on MMRM model
ENDPOINTS BY VISIT THROUGH WEEK 24 AT WEEK-24 DBL, AFTER WEEK 24 THROUGH WEEK 52 AT WEEK-52 DBL, AND AFTER WEEK 52 THROUGH WEEK 112 AT WEEK-112 DBL USING TREATMENT POLICY ESTIMAND AND THE FAS-UKR ANALYSIS SET INSTEAD OF FAS				
The endpoints analyzed using the Adjusted Composite Estimand for corresponding visits/DBL are repeated here, except: • No imputation will be applied, and all analyses will be based on observed data. • Continuous endpoints will use ANCOVA instead of MMRM; binary response endpoints will use CMH test with CI based on Wald statistics instead of GLMM for visits up to and including Week 24.				

5.3.6.4. Endpoints Related to Health-Related Quality of Life and Other Outcomes

In this study, Other Endpoints for HRQOL measures include questionnaires of CCI [redacted]

All Other Endpoints related to CCI [redacted]

5.3.6.4.1. CCI Related Endpoints

CCI [redacted]

[redacted]

CCI [redacted]

[redacted]

The CCI related Other Endpoints include:

CCI [redacted]

CC [redacted]

5.3.6.4.2. CCI [redacted]

CCI [redacted]

The CCI [redacted] related Other Endpoints include:

CC [redacted]

5.3.6.4.3. CCI [redacted] Related Endpoints

CCI [redacted]

[redacted]

CCI [redacted]

The CCI [redacted] related Other Endpoints include:

[redacted]

5.3.6.4.4. PsAID-12 Related Endpoints

[redacted]

The [redacted] related Other Endpoints include:

[redacted]

5.3.6.4.5. Method of Analysis

In general, the same methods of analysis for the Other Efficacy endpoints related to reduction of signs and symptoms and physical function in Section 5.3.6.2.9 will also be used for those related to HRQOL.

Table 10 summarizes the analyses for supportive efficacy endpoints related to HRQOL, the methods for analyses, and the data handling rules used.

Table 10: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Endpoints of HRQOL

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
ENDPOINTS BY VISIT THROUGH WEEK 24 AT WEEK-24 DBL, ADJUSTED COMPOSITE ESTIMAND				
1	Change from baseline in SF-36 PCS/MCS score	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
2	Change from baseline in SF-36 domain scales score	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
3	Change from baseline in PROMIS-29 score	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
4	Change from baseline in FACIT-F score	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
5	Proportions of participants with ≥4-point improvement from baseline in FACIT-F score	FAS	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates, and treatment difference in response rates and 95% CI and p-value, based on GLMM
6	Change from baseline in PsAID-12 score	FAS	-	• Summarized descriptively • LS mean (95% CI) for each intervention group, LS mean difference (95% CI) and p-values for differences between groups based on MMRM
'-' indicates no missing data rules to be applied				
ENDPOINTS BY VISIT AFTER WEEK 24 THROUGH WEEK 52 AT WEEK-52 DBL, USING ADJUSTED COMPOSITE ESTIMAND				
1, 2, 3, 4, 6	Same as through Week 24	Same as through Week 24	Same as through Week 24	• Summarized descriptively • LSMeans and 95% CI by intervention group based on MMRM model
5	Same as through Week 24	Same as through Week 24	Same as through Week 24	• Summarized descriptively • Response rates by intervention group based on GLMM
ENDPOINTS BY VISIT AFTER WEEK 52 THROUGH WEEK 112 AT WEEK-112 DBL, USING ADJUSTED COMPOSITE ESTIMAND				

Table 10: Summary of Estimands, Analysis Sets, Data Handling Rules, and Analysis Methods for Endpoints of HRQOL

	Endpoint	Analysis Set	Missing Data Rules	Analysis Methods
1, 2, 3, 4, 6	Same as through Week 24	Same as through Week 24	Same as through Week 24	• Summarized descriptively • LSMeans and 95% CI by intervention group based on MMRM model
5	Same as through Week 24	Same as through Week 24	NRI for MAR not due to Natural Disaster or Major Disruption	• Summarized descriptively • Response rates by intervention group based on GLMM
ENDPOINTS BY VISIT THROUGH WEEK 24 AT WEEK-24 DBL, AFTER WEEK 24 THROUGH WEEK 52 AT WEEK-52 DBL, AND AFTER WEEK 52 THROUGH WEEK 112 AT WEEK-112 DBL USING TREATMENT POLICY ESTIMAND AND THE FAS-UKR ANALYSIS INSTEAD OF FAS				
The endpoints analyzed using the Adjusted Composite Estimand for corresponding visits/DBL are repeated here, except: • No imputation will be applied, and all analyses will be based on observed data. • Continuous endpoints will use ANCOVA instead of MMRM; binary response endpoints will use CMH test with CI based on Wald statistics instead of GLMM for visits up to and including Week 24.				

5.4. Safety Analyses

All safety analyses will be based on the safety analysis set based on **actual** intervention received. Safety will be assessed by summarizing the occurrences and type of AEs, vital signs (pulse, blood pressure, and weight) and examining the changes in the laboratory parameters. No formal statistical comparison is planned.

5.4.1. Safety Tables Presentation

There are 3 DBLs in this study, respectively, at Week 24, Week 52, and End of Study (Week 112). Depending on the safety data categories, the cumulative safety data will be analyzed through different study periods which include, but are not limited to, through Week 24 and through end of study periods. Tabular summaries of safety events for key study periods are in general presented as follows:

5.4.1.1. Summaries Through Week 24

Safety data through Week 24 will be analyzed according to the following intervention groups:

1. **Placebo**: Participants who received placebo only and no guselkumab prior to Week 24.
2. **Guselkumab 100 mg at Weeks 0, 4, and then q8w**: Participants who received guselkumab 100 mg q8w prior to Week 24 with an additional dose at Week 4.
3. **Guselkumab 100 mg q4w**: Participants who received guselkumab 100 mg q4w prior to Week 24.
4. **Guselkumab Combined:** Participants in Groups 2 and 3.

The above intervention groups 1-3 are **mutually exclusive**. This allows between-group comparisons of safety between a guselkumab group and the placebo group based on similar follow-up period in each group. The safety tables will have the column headings below:

		Guselkumab		
	Placebo	100 mg q8w	100 mg q4w	Combined
Analysis set: Safety Analysis Set	###	###	###	###

For participants who started treatment with placebo only but later received any amount of guselkumab prior to Week 24 inadvertently, the safety events/measurements on and after the first dose of guselkumab, will be excluded from the data summaries through Week 24. Only the safety events/measurements that occurred while the participants had been receiving placebo only will be included in the data summaries through Week 24.

5.4.1.2. Summaries Through Week 52

Safety data through Week 52 will be analyzed according to the following intervention groups:

1. **Placebo:** Participants who received placebo only. Follow-up will be based on the period that the participant was on placebo from the first dose up to Week 52.
    a. For participants who started treatment with placebo and later received treatment with guselkumab (due to CO or inadvertently), follow-up will end at the first dose of guselkumab, and only the safety events/measurements that occurred prior to the first dose of guselkumab will be included in this group.
2. **Placebo → Guselkumab 100 mg q4w:** Participants who started treatment with placebo and later received treatment with guselkumab (due to CO or inadvertently).Follow-up will start from the first dose of guselkumab up to end of study. All the safety events/measurements that occurred on and after the first dose of guselkumab up to end of study will be included in this group.
3. **Guselkumab 100 mg at Weeks 0, 4, and then q8w**: Participants who received guselkumab 100 mg q8w prior to Week 24 with an additional dose at Week 4. Follow-up will be from the first dose up to Week 52.
4. **Guselkumab 100 mg q4w**: Participants who received guselkumab 100 mg q4w prior to Week 24. Follow-up will be from the first dose up to end of study.
5. **Guselkumab 100 mg q4w Combined:** Participants in Groups 2 and 4.
6. **All Guselkumab Combined:** Participants in Groups 2, 3, and 4.

The above intervention groups 1-2 are **not mutually exclusive**. The safety tables will have the column headings below:

		Guselkumab				
	Placebo	Placebo → 100 mg q4w	100 mg q8w	100 mg q4w	100 mg q4w Combined	All Combined
Analysis set: Safety Analysis Set	###	###	###	###	###	###

5.4.1.3. Summaries Through End of Study

Safety data through End of Study (Week 112) will be analyzed similarly to safety data through Week 52, replacing Week 52 with End of Study.

5.4.2. Extent of Exposure and Study Follow-up

The number and percentage of participants who receive study intervention will be summarized. Descriptive statistics for duration study intervention (N, mean, SD, median, and range (minimum, maximum)) will be summarized.

Study intervention duration is defined as (date of last dose of study intervention – date of first dose of study intervention) +1. For the placebo intervention group which has planned crossover at Week 24, the study intervention duration prior to first guselkumab dose will be summarized separately to the study intervention duration on/after first guselkumab dose.

Study follow-up duration is defined in Section 5.4.1.

Study intervention compliance will be summarized descriptively. See Section 6.6 for further details.

5.4.3. Adverse Events

The verbatim terms used in the eCRF by investigators to identify AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE occurring at or after the initial administration of study intervention and those AEs that were present at baseline but worsened in severity after the start of initial study intervention are considered to be treatment emergent. If the event occurs on the day of the initial administration of study intervention, and either event time or time of administration are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered to be treatment emergent unless it is known to be prior to the first administration of study intervention based on partial onset date or resolution date. All reported treatment-emergent adverse events will be included in the analysis. For each adverse event, the number and percentage of participants who experience at least 1 occurrence of the given event will be summarized by intervention group.

Summary tables will be provided for treatment-emergent adverse events:

- Treatment emergent adverse events (TEAEs)
- Treatment emergent SAEs
- TEAEs with severe intensity
- TEAEs leading to permanent discontinuation of study intervention
- TEAEs reasonably related to study intervention
- SAEs reasonably related to study intervention

- TEAEs leading to dose interruption of study intervention
- Treatment emergent infections
- Treatment emergent serious infections
- Treatment emergent infections requiring oral or parenteral anti-microbial treatment
- Injection-site reactions
- Venous thromboembolism (VTE) events
- Clinically important hepatic disorder events
- Anaphylaxis, hypersensitivity and serum sickness reactions
- TEAEs leading to death

All AE summary tables will include average weeks of follow-up and average number of study intervention for each intervention group.

In addition to the summary tables, listings will be provided for participants who had:

- SAEs
- TEAEs leading to permanent discontinuation of study intervention
- Anaphylactic reactions or serum sickness reactions
- Malignancies
- Serious infections including TB
- TEAEs leading to death
- Venous thromboembolism (VTE) events
- Major adverse cardiovascular events (MACE) events
- Clinically important hepatic disorder events
- Opportunistic infections

Section 6.7 contains the methods of identification for selected AEs of interest.

A listing of participants who died will be provided, including cause of death, date of death, relationship to study intervention, and study day of death in relation to reference date.

5.4.4. Additional Safety Assessments

5.4.4.1. Clinical Laboratory Tests

Clinical laboratory tests will be displayed for the participants included in the safety analysis set.

The clinical laboratory parameters to be evaluated by the central laboratory include but are not limited to:

- **Hematology**: basophils, eosinophils, hemoglobin, hematocrit, lymphocytes, monocytes, neutrophils, platelets, red blood cell (RBC) count and white blood cell (WBC) count

- **Clinical chemistry**: albumin, alkaline phosphatase, alanine aminotransferase (serum glutamate pyruvate transaminase) [ALT (SGPT)], aspartate aminotransferase (serum glutamic oxaloacetic transaminase) [AST (SGOT)], bicarbonate, blood urea nitrogen (BUN), calcium, chloride, creatinine, glucose, potassium, sodium, total bilirubin, and total protein

Due to the Major Disruption resulting from the conflict involving Ukraine and neighboring countries/territories beginning 24 February 2022, central labs were unavailable for some sites over a period of time. In these instances, local labs, limited to the following parameters, may have been entered into the eCRFs: sodium, potassium, chloride, bicarbonate, BUN, creatinine, total bilirubin, direct bilirubin (conditionally), indirect bilirubin (conditionally), AST, ALT, gamma-glutamyltransferase (GGT), alkaline phosphatase (ALP), and LDH. Certain analyses will use central lab data only, while other analyses may use a combination of central lab data and local lab data.

Descriptive statistics and graphical displays of observed values and changes from baseline will be presented for selected chemistry and hematology laboratory tests at scheduled time points. Only central lab data will be used for this analysis.

Shift tables from baseline to post-baseline will be produced for select laboratory parameters. Both central and local lab data will be used for this analysis.

Abnormality criteria based on toxicity grade will be applied to baseline and post-baseline values using National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI-CTCAE) for parameters with NCI-CTCAE criteria defined (Appendix 9). Applicable laboratory results will be graded according to NCI-CTCAE version 5.0. Both central and local lab data will be used for this analysis. Abnormality for selected chemistry and hematology laboratory tests will be summarized by study intervention for:

- Number and percent of participants with post-baseline values by maximum toxicity grade
- Listings of participants with any post-baseline lab value of NCI-CTCAE toxicity Grade 3 or higher

Number and percent of participants with post-baseline elevated liver chemistry tests will also be produced using both central and local lab data, for:

ALT categories

- >1x to <3x Upper limit of normal (ULN)
- ≥3× to <5×ULN
- ≥5× to <8×ULN
- ≥8×ULN

AST categories

- >1× to <3×ULN
- ≥3× to <5×ULN

- ≥5× to <8×ULN
- ≥8×ULN

Total Bilirubin categories:

- >1× to <2×ULN
- ≥2×ULN

ALP categories:

- >1× to <2×ULN
- ≥2× to <4×ULN
- ≥4×ULN

Both central and local lab data will be used for this analysis.

Number and percent of participants who met each of the following 5 liver function biochemical criteria individually, as well as overall (ie, met any of the 5), as determined by the investigator, will be summarized:

- **ALT or AST absolute**:

  ALT or AST ≥8×ULN

- **ALT or AST increase**:

  ALT or AST ≥5×ULN but <8×ULN persists for ≥2 weeks, OR

  ALT or AST ≥3×ULN but <5×ULN persists for ≥4 weeks

- **ALT/AST/Bilirubin or ALT/AST/INR:**

  ALT or AST ≥3×ULN **and** total bilirubin ≥2×ULN, OR

  ALT or AST ≥3×ULN **and** INR >1.5

- **Cannot monitor**:

  ALT or AST ≥5×ULN but <8×ULN and cannot be monitored weekly for ≥2 weeks, OR

  ALT or AST ≥3×ULN but <5×ULN and cannot be monitored weekly for ≥4 weeks

- **Symptomatic**:

  ALT or AST ≥3×ULN associated with symptoms (new or worsening) believed to be related to liver injury or hypersensitivity

Listings of participants meeting ANY biochemical criteria of post-baseline ALT≥3x ULN, AST≥3x ULN, ALP≥2x ULN, OR total bilirubin≥2x ULN will be created. Additionally, participants will be assessed for the following 2 Combined Biochemical Criteria:

1) Total bilirubin ≥2xULN within 5 days after either ALT or AST ≥3xULN.
2) INR >1.5 within 5 days after either ALT or AST ≥3xULN.

5.4.4.2. Vital Signs and Physical Examination Findings

Continuous vital sign parameters including pulse, blood pressure (systolic and diastolic), and weight will be summarized at each assessment time point. The observed value and change from baseline will be summarized by intervention group. Descriptive statistics (mean, SD, median, minimum, and maximum) will be presented.

Incidence of markedly abnormal vital signs during intervention, as defined in Table 11, will be summarized for participants who had a baseline assessment and at least 1 post-baseline assessment for that vital sign. A listing of participants with markedly abnormal vital signs will be presented.

Table 11: Markedly Abnormal Vital Signs

Vital Sign	Criteria
Pulse	>120 bpm and with >30 bpm increase from baseline
	<50 bpm and with >20 bpm decrease from baseline
Systolic blood pressure	>180 mm Hg and with >40 mm Hg increase from baseline
	<90 mm Hg and with >30 mm Hg decrease from baseline
Diastolic blood pressure	>105 mm Hg and with >30 mm Hg increase from baseline
	<50 mm Hg and with >20 mm Hg decrease from baseline

5.4.4.3. Electrocardiogram

No analysis is planned.

5.4.4.4. Other Safety Parameters

5.4.4.4.1. Suicidal Ideation and Behavior

The electronic Columbia-Suicide Severity Rating Scale (eC-SSRS) will be used as a screening tool to prospectively evaluate the potential of guselkumab to induce suicidal ideation and behavior. The eC-SSRS defines 5 subtypes of suicidal ideation and behavior in addition to self-injurious behavior with no suicidal intent, and is a fully-structured participant self-report questionnaire, including standardized questions, follow-up prompts, error handling routines, and scoring conventions (Mundt 2013; Posner 2011). Two versions of the eC-SSRS will be used in this study, the Lifetime version and the Since Last Contact version. The Lifetime version will be conducted during the screening visit and the Since Last Contact version will be conducted at all other visits through Week 112.

Participants will complete the eC-SSRS questionnaire using the Sponsor-provided electronic tablets (or through an Interactive Voice Response System, if available). Study site personnel will train the participants on how to use the electronic device and/or a telephone system. The eC-SSRS will be provided in the local languages in accordance with local guidelines.

The eC-SSRS will be performed during each evaluation visit according to the Time and Events schedule. The eC-SSRS should be performed after the joint assessment at the screening visit (after signing informed consent). At Week 0/baseline and at all post-baseline visits, the eC-SSRS will be the first assessment/questionnaire that the participant completes prior to study intervention administration.

At the conclusion of each assessment, the site will receive an eC-SSRS Findings Report from the eC-SSRS vendor. Positive reports are generated from the eC-SSRS vendor for ANY of the following findings:

- Ideation Level 4: Some intent to act, no plan
- Ideation Level 5: Specific plan and intent
- Behaviors: Actual Suicide Attempts
- Behaviors: Interrupted Attempts
- Behaviors: Aborted Attempts
- Behaviors: Preparatory actions

Negative suicidality indication reports are generated from the eC-SSRS vendor when there are NO indications of the above.

Any eC-SSRS findings, which in the opinion of the investigator are new or considered to be worsening and clinically significant, should be reported on the AE eCRF.

Suicidal ideation and behavior will be categorized as follows, with higher scores indicating greater severity:

Suicidal Ideation (1-5)

1 = Wish to be Dead

2 = Non-specific Active Suicidal Thoughts

3 = Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act

4 = Active Suicidal Ideation with Some Intent to Act, without Specific Plan

5 = Active Suicidal Ideation with Specific Plan and Intent

Suicidal Behavior (6-10)

6 = Preparatory Acts or Behavior

7 = Aborted Attempt

8 = Interrupted Attempt

9 = Actual Attempt (non-fatal)

10 = Completed Suicide

The baseline is defined as the most severe/maximum score at screening and Week 0. Suicidal ideation and behavior will be analyzed by the most severe/maximum post-baseline outcome. In addition, a shift table from baseline to post-baseline will also be provided. Participants with positive (ie, score >0) ideation and behavior will be presented in a data listing.

5.5. Other Analyses

5.5.1. Pharmacokinetics

PK analyses will be performed on the PK analysis set, defined as participants who have received at least 1 complete dose of guselkumab and have at least 1 valid blood sample drawn for PK analysis (Section 4.6). Subjects will be analyzed according to the treatment groups that they actually receive. No imputation for missing concentration data will be performed.

Descriptive statistics (N, mean, SD, median, range, CV (%) and IQ range) will be used to summarize serum guselkumab concentrations at each sampling time point by treatment group. PK data may be displayed graphically. The following analyses will be performed by treatment group as appropriate:

- Summary of serum guselkumab concentrations at each visit by treatment group
- Proportion of subjects without detectable serum guselkumab concentration at each visit by treatment group
- Summary of serum guselkumab concentrations at each visit by treatment group and body weight
- Summary of serum guselkumab concentrations at each visit by treatment group and baseline MTX use (Yes, No)
- Summary of serum guselkumab concentrations by baseline CRP levels
- Plot of median (IQ) serum guselkumab concentrations over time by treatment group

In addition, the relationship between serum guselkumab concentrations and safety or efficacy may be explored.

For summary statistics of serum guselkumab concentrations, concentration values below the LLOQ will be treated as zero. Once a subject meets one of the following dosing deviation criteria, the subject's data will be excluded from the by-visit data analyses from that point onwards.

Dosing deviation criteria:

- Discontinue SC guselkumab administrations.
- Skipped an SC guselkumab administration.
- Received an incomplete/ incorrect SC dose.
- Received an incorrect SC study agent.
- Received an additional SC guselkumab dose.

In addition, if a subject has an administration outside of visit windows (Section 5.1.1), the concentration data collected at and after that visit will be excluded from the by-visit data analyses. Additional exclusions for incongruous PK data to be implemented based on Janssen SOP-07952.

5.5.2. Immunogenicity

The antibodies to guselkumab will be summarized based on all participants who received at least one (complete or partial) administration of guselkumab and who had at least 1 sample obtained after their first administration of guselkumab (Section 4.7). Subjects will be analyzed according to the treatment groups that they actually receive. No imputation for missing concentration data will be performed.

The following analysis of antibodies to guselkumab will be performed by treatment group:

- Summary of antibodies to guselkumab status
- Summary of neutralizing antibodies to guselkumab status
- List of subjects positive for antibodies to guselkumab

In addition, to explore the relationship between antibodies to guselkumab status and serum guselkumab concentrations, efficacy and safety, the following analysis may be performed as appropriate:

- Summary of clinical response (eg, ACR 20 and ACR50, IGA) by antibody to guselkumab Status
- Summary of injection-site reactions by antibody to guselkumab status
- Summary of serum guselkumab concentrations by antibody to guselkumab status
- Plots of median (IQ) trough serum guselkumab concentrations over time by antibody to guselkumab status.

5.5.3. Biomarker/Pharmacodynamic Analysis

Methods and results for biomarker/PD analyses will be presented in a separate technical report.

5.5.4. Pharmacokinetic/Pharmacodynamic Relationships

If data permit, the relationships between serum guselkumab concentration and efficacy may be analyzed graphically. If any visual trend is observed, a suitable population PK/PD model may be developed to describe the exposure-response relationship. Details will be given in a population PK/PD analysis plan and results of the population PK/PD analysis will be presented in a separate technical report.

5.5.5. Definition of Subgroups

To evaluate the consistency in the primary efficacy endpoint (proportion of participants who achieve ACR 20 at Week 24) and the major secondary endpoint over demographics, baseline characteristics, prior and baseline medication use, subgroup analyses will be performed. The subgroups include, but are not limited to, the following:

Subgroup	Variant	Definition
Demographics		
Gender		• Male • Female
Race		• White • Other
Age at baseline (year)		• <45 • ≥45 and <65 • ≥65
Body weight at baseline (kg)	1 (Categories)	• ≤90 • >90
Body weight at baseline (kg)	2 (Quartiles)	• 1st Quartile: ([##] to [##]) • 2nd Quartile: ([##] to [##]) • 3rd Quartile: ([##] to [##]) • 4th Quartile: ([##] to [##])
Body mass index at baseline (kg/m^2)		• Normal [<25] • Overweight [≥25 to <30] • Obese [≥30]
Participating regions		• Eastern Europe (except Russia, Ukraine, Poland) • North America • Poland • Russia • Ukraine • Other countries
Baseline disease characteristics		
PsA duration at baseline (year)		• <1 • ≥1 to <3 • ≥3
PsA subtype		• distal interphalangeal joint involvement • polyarticular arthritis with absence of rheumatoid nodules • asymmetric peripheral arthritis • spondylitis with peripheral arthritis
Number of swollen joints at baseline		• <10 • 10 to 15 • >15
Number of tender joints at baseline		• <10 • 10 to 15 • >15
HAQ-DI score at baseline		• <1 • 1 to 2 • >2
CRP at baseline (mg/dL)	1 (Categories)	• <1 • 1-2 • ≥2
CRP at baseline (mg/dL)	2 (Quartiles)	• 1st Quartile: ([##] to [##]) • 2nd Quartile: ([##] to [##]) • 3rd Quartile: ([##] to [##]) • 4th Quartile: ([##] to [##])
Dactylitis at baseline		• Yes • No
Enthesitis at baseline		• Yes • No

Subgroup	Variant	Definition
PASI score at baseline		• <12 • ≥12 to <20 • ≥20
BSA of psoriasis at baseline		• <3% • ≥3% to <10% • ≥10% to <20% • ≥20%
IGA score at baseline		• <2 • ≥2
Prior and baseline medication use		
Reason for discontinuation for prior anti-TNF		• Efficacy - inadequate response (IR) • Safety - contraindication or intolerance (but not IR) • Other
Use of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) at baseline (based on eCRF)		• Yes • No
Oral corticosteroids at baseline (based on eCRF)		• Yes • No
NSAIDs at baseline		• Yes • No
Number of prior non-biologic treatments including DMARDs, systemic immunosuppressive drugs, and apremilast		• 0 • 1 • 2 • ≥3
Reason for discontinuation of prior DMARDs		• Efficacy - inadequate response (IR) • Safety - contraindication or intolerance (but not IR) • Other
Note that some of the above subgroup cut-off points may be changed if there are no or few participants within a subgroup category		

5.6. Interim Analyses

No interim analysis is planned for this study.

5.6.1. Data Monitoring Committee (DMC)

No DMC is planned for this study.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1: List of Abbreviations

ACR	American College of Rheumatology
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	Analysis of Covariance
AST	aspartate aminotransferase
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BMI	body mass index
BSA	body surface area
BUN	blood urea nitrogen
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRP	C-reactive protein
CTCAE	Common Terminology Criteria for Adverse Events
DAPSA	Disease Activity index for PSoriatic Arthritis
DBL	database lock
DMARD	Disease-Modifying Antirheumatic Drugs
eCRF	electronic case report form
eC-SSRS	electronic Columbia-Suicide Severity Rating Scale
EE	early escape
FACIT-F	Functional Assessment of Chronic Illness Therapy - fatigue
FAS	Full Analysis Set
FAS-UKR	Full Analysis Set Excluding Ukraine
FCS	Full Conditional Specifications
GDEV	Physician's Global Assessment of Disease Activity
GDPT	Patient's Global Assessment of Disease Activity
GGT	gamma-glutamyltransferase
GLMM	Generalized Linear Mixed Model
HAQ-DI	HAQ disability index
HCQ	hydroxychloroquine
HRQOL	health related quality of life
ICE	intercurrent event
IGA	Investigator's Global Assessment
INR	International normalized ratio
IQ	interquartile
IR	inadequate response
IWRS	interactive web response system
LEF	leflunomide
LEI	Leeds Enthesitis Index
LLN	lower limit of normal
LLOQ	lower limit of quantification
LSmeans	least squares means
MACE	Major adverse cardiovascular events
MAR	missing at random
MCS	Mental Component Score
MDA	Minimal Disease Activity
MI	multiple imputation
MMRM	Mixed-Effect Model Repeated Measures
MNAR	Missing not at random
MTX	methotrexate
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NR	non-responder
NRS	numeric rating scale
NSAID	non-steroidal anti-inflammatory drug

PAIN	Patient's assessment of pain
PASDAS	Psoriatic ArthritiS Disease Activity Score
PASI	Psoriasis Area and Severity Index
PCS	Physical Component Summary
PD	pharmacodynamic(s)
PK	pharmacokinetic(s)
PPAS	Per-Protocol Analysis Set
PROMIS-29	Patient-Reported Outcomes Measurement Information System-29
PsA	psoriatic arthritis
PsAID-12	PsA Impact of Disease-12
q4w	every 4 weeks
q8w	every 8 weeks
RA	rheumatoid arthritis
SAE	serious adverse event
SAP	Statistical Analysis Plan
SC	subcutaneous
SD	standard deviation
SDV	source data verification
SF-36	36-Item Short Form Survey Instrument
SJC	Swollen Joint Count
SPARCC	Spondyloarthritis Research Consortium of Canada
SSZ	sulfasalazine
TEAE	treatment emergent adverse event
TF	treatment failure(s)
TJC	Tender Joint Count
TNF	tumor necrosis factor
ULN	upper limit of normal
VAS	visual analog scale
VLDA	Very Low Disease Activity
VTE	Venous thromboembolism

6.2. Appendix 2: Demographics and Baseline Characteristics

Table 12 presents a list of the demographic and baseline variables that will be summarized by intervention group, combined active intervention group, and overall, for the following analysis sets: FAS.

Table 12: Demographic Variables

Continuous Variables:	**Summary Type**
Age (years)	Descriptive statistics (N, mean, SD, median and range [minimum and maximum], and IQ range).
Weight (kg)	
Height (cm)	
Body Mass Index (BMI) (kg/m^2)	
Categorical Variables	
Age (45 years, 45-<65 years, >=65 years)	Frequency distribution with the number and percentage of participants in each category.
Sex (male, female, undifferentiated)	
Race[a] (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)	
Weight (≤90kg, >90kg)	
BMI (underweight <18.5 kg/m^2, normal 18.5-<25 kg/m^2, overweight 25-<30 kg/m^2, obese >=30 kg/m^2)	

[a]If multiple race categories are indicated, the Race is recorded as 'Multiple'

The baseline characteristics will be summarized for the same analysis sets as the demographic variables. They include, but are not limited to, baseline disease characteristics of PsA (eg, duration of disease, PsA subtypes, baseline efficacy assessments), medical history (Appendix 5), prior exposure to non-biologic medications, prior joint procedures/injections, and baseline medication usage for PsA (Appendix 4).

6.3. Appendix 3: Protocol Deviations

In general, the following list of major protocol deviations may have the potential to impact participants' rights, safety or well-being, or the integrity and/or result of the clinical study. Participants with major protocol deviations will be identified prior to database lock and the participants with major protocol deviations will be summarized by category.

- Developed withdrawal criteria but not withdrawn
- Entered but did not satisfy criteria
- Received a disallowed concomitant treatment
- Received wrong treatment or incorrect dose
- Other

The study selection criteria will be grouped into the following 5 categories: PsA disease criteria, medication criteria, laboratory criteria, medical history criteria, and other.

Protocol deviation in study intervention administrations includes missing doses, incorrect doses, and treatments administered out of the dosing windows defined in Section 5.1.1.1. Additionally, missed doses due to Major Disruption or Natural Disaster will be summarized.

6.4. Appendix 4: Prior and Concomitant Medications

Prior and Concomitant medications will be coded using the [World Health Organization Drug Dictionary (WHO-DD)]. Prior medications are defined as any therapy used before the day of first dose (partial or complete) of study intervention. Concomitant medications are defined as any therapy used on or after the same day as the first dose of study intervention, including those that started before and continue on after the first dose of study intervention.

Summaries of concomitant medications will be presented by ATC class and ATC term, intervention group. The proportion of participants who receive each concomitant medication will be summarized as well as the proportion of participants who receive at least 1 concomitant medication. In addition, concomitant medications of special interest will be presented. These include non-biologic DMARDs, systemic corticosteroids, and NSAIDs. See Section 6.8 for list of medications in each category.

Prior medications taken for PsA and/or psoriasis (eg, non-biologic DMARDs, apremilast, immunosuppressives, and NSAIDS) will be summarized by randomized intervention group.

6.5. Appendix 5: Medical History

Number and percentage of participants who had medical histories of interest for PsA will be collected and summarized by intervention group, including:

- Degenerative disk disease
- Gout
- Fibromyalgia
- Inflammatory Back Pain
- Osteoarthritis
- Angina Pectoris
- Coronary Artery Bypass Graft
- Coronary Artery Disease
- Diabetes Mellitus
- Family history of early coronary artery disease less than 55
- Hyperlipidemia
- Hypertension
- Myocardial Infarction
- Percutaneous Coronary Intervention
- Peripheral Vascular Disease
- Stroke
- Transient Ischemic Attack
- Asthma
- Chronic Liver Disease (eg fatty liver, alcohol-induced, cirrhosis)
- Hidradenitis Suppurativa
- Inflammatory Bowel Disease
- Uveitis
- Squamous Cell Carcinoma
- Basal Cell carcinoma
- Other malignancy
- Family history of cancer in a 1st degree relative (not Squamous Cell or Basal Cell Carcinoma)
- Anxiety
- Depression
- COVID-19

- Hospitalized within the past year (excluding pregnancy)
- Hospitalized within the past year for infection

6.6. Appendix 6: Intervention Compliance

Compliance will be summarized descriptively for the overall study intervention, as well as for guselkumab and placebo separately. Compliance to randomized intervention versus actual intervention will be presented in a summary table and will be calculated as (the number of injections completed / the number of injections planned * 100).

6.7. Appendix 7: Adverse Events of Special Interest

Adverse events of special interest, as well as other adverse events of interest, will be identified based on criteria specified in the following table.

Type of Adverse Event	**MedDRA Terms Search Methodology**	**Requires Medical Review**
Adverse Events of Special Interest		
Malignancy	Malignant tumors (SMQ-narrow scope).	Yes
Active Tuberculosis	HLT of Tuberculosis infections excluding PT of Latent Tuberculosis	Yes
Other Adverse Events of Interest		
Infections	SOC Infections and infestations Note that serious infections, and infections requiring oral or parenteral anti-microbial treatment, are based on this MedDRA determination as well as eCRF checkboxes for serious AE and for requiring oral or parenteral anti-microbial treatment respectively	No
Opportunistic Infections	Opportunistic infections (SMQ-narrow scope)	Yes
ISR	No MedDRA search used. Based completely on eCRF checkbox	No
Anaphylaxis and Hypersensitivity	PTs of Anaphylactic reaction, Anaphylactic shock, Anaphylactoid reaction, Anaphylactoid shock, and Type I Hypersensitivity	No
Serum Sickness Reactions	PTs of serum sickness and Serum sickness-like reaction	No
MACE	**SMQs:** • Myocardial infarction (narrow) • Ischaemic central nervous system vascular conditions (narrow scope) • Haemorrhagic central nervous system vascular conditions (narrow scope) • **PTs:** Sudden death **AESOC:**	Yes

	• Cardiac Disorders (fatal events only) • Vascular Disorders (fatal events only)	
VTEs	Customized MedDRA PTs related to venous thrombosis and embolism involving the deep venous vasculature: Axillary vein thrombosis, Brachiocephalic vein thrombosis, Budd-Chiari syndrome, Deep vein thrombosis, Deep vein thrombosis postoperative, Embolism venous, Hepatic vein thrombosis, Homans' sign positive, Inferior vena cava syndrome, Jugular vein thrombosis, Mahler sign, May-Thurner syndrome, Mesenteric vein thrombosis, Obstetrical pulmonary embolism, Ovarian vein thrombosis, Paget-Schroetter syndrome, Pelvic venous thrombosis, Penile vein thrombosis, Peripheral vein thrombus extension, Peripheral vein thrombosis, Post procedural pulmonary embolism, Postpartum venous thrombosis, Pulmonary embolism, Pulmonary infarction, Pulmonary microemboli, Pulmonary thrombosis, Pulmonary venous thrombosis, Renal vein thrombosis, Spermatic vein thrombosis, Splenic vein thrombosis, Subclavian vein embolism, Subclavian vein thrombosis, Thrombosis corpora cavernosa, Vena cava embolism, Vena cava thrombosis, Venous thrombosis, Venous thrombosis in pregnancy, Venous thrombosis limb, Visceral venous thrombosis	No
Clinically Important Hepatic Disorders	Drug related hepatic disorders - comprehensive search (SMQ – narrow scope) and either SAE or AE leading to discontinuation of study intervention	No

6.8. Appendix 8: Medications of Special Interest

Concomitant medications of special interest are defined as follows:

Concomitant Medication Special Interest Category	**Categories**
Other biologic treatments or other experimental treatments	
Non-biologic DMARD	MTX, SSZ, LEF, HCQ, chloroquine, gold preparations, penicillamine, other non-biologic DMARDs
Oral corticosteroids	Oral corticosteroids
NSAIDs	NSAIDs

Prior Medication Special Interest Category	**Categories**
Prior anti-TNF agents	
Non-biologic DMARD	MTX, SSZ, LEF, HCQ, chloroquine, gold preparations, penicillamine, other non-biologic DMARDs
Immunosuppressives	Cyclosporine, azathioprine, mycophenolate mofetil, tacrolimus, other immunosuppressives
Systemic corticosteroids	Systemic corticosteroids
NSAIDs	NSAIDs
Apremilast	Apremilast
Previous medications and therapies for PsA	Non-biologic DMARDs, immunosuppressives, apremilast, systemic corticosteroids, NSAIDs
Previous medications and therapies for PsO	Cyclosporine, Topical, Acitretin, UVB, Apremilast, PUVA

6.9. Appendix 9: Laboratory Toxicity Grading

The grading scale use for lab assessments is based on CTCAE v5.0.

If a laboratory value falls within the grading as specified below but also within the local laboratory normal limits, the value is considered to be normal and will be reset to grade 0.

Pre-baseline measurements will use the same grading ranges as applied to baseline measurements. In case a test has two sets of ranges – one for baseline normal and one for baseline abnormal, the one for baseline normal will be applied for all measurements taken pre-baseline and on baseline.

Text in gray italic in the table is present in the grading scale but is not applied by Janssen when grading lab data.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Blood and lymphatic system disorders					
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hemoglobin (Hgb) <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hemoglobin (Hgb) <8.0 g/dL; <4.9 mmol/L; <80 g/L; *transfusion indicated*	*Life-threatening consequences; urgent intervention indicated*	Clinical signs and symptoms are not taken into consideration for grading.
Leukocytosis	-	-	>100,000/mm^3; >100 x $10e^9$ /L	*Clinical manifestations of leucostasis; urgent intervention indicated*	Clinical signs and symptoms are not taken into consideration for grading; Added ranges in SI unit (x $10e^9$ /L).
Investigations					
Activated partial thromboplastin time prolonged	>ULN - 1.5×ULN	>1.5 - 2.5×ULN	>2.5×ULN; *bleeding*	-	Clinical signs and symptoms are not taken into consideration for grading.
Alanine aminotransferase increased	>ULN - 3.0×ULN if baseline was normal; 1.5 - 3.0×baseline if baseline was abnormal	>3.0 - 5.0×ULN if baseline was normal; >3.0 - 5.0×baseline if baseline was abnormal	>5.0 - 20.0×ULN if baseline was normal; >5.0 - 20.0×baseline if baseline was abnormal	>20.0×ULN if baseline was normal; >20.0×baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline >ULN. If baseline <LLN, then ranges for "normal baseline" are applied.
Alkaline phosphatase increased	>ULN - 2.5×ULN if baseline was normal; 2.0 - 2.5×baseline if baseline was abnormal	>2.5 - 5.0×ULN if baseline was normal; >2.5 - 5.0×baseline if baseline was abnormal	>5.0 - 20.0×ULN if baseline was normal; >5.0 - 20.0×baseline if baseline was abnormal	>20.0×ULN if baseline was normal; >20.0×baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline >ULN. If baseline <LLN, then ranges for "normal baseline" are applied.
Aspartate aminotransferase increased	>ULN - 3.0×ULN if baseline was normal; 1.5 - 3.0×baseline if baseline was abnormal	>3.0 - 5.0×ULN if baseline was normal; >3.0 - 5.0×baseline if baseline was abnormal	>5.0 - 20.0×ULN if baseline was normal; >5.0 - 20.0×baseline if baseline was abnormal	>20.0×ULN if baseline was normal; >20.0×baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline >ULN. If baseline <LLN, then ranges for

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
					"normal baseline" are applied.
Blood bilirubin increased	>ULN - 1.5×ULN if baseline was normal; >1.0 - 1.5×baseline if baseline was abnormal	>1.5 - 3.0×ULN if baseline was normal; >1.5 - 3.0×baseline if baseline was abnormal	>3.0 - 10.0×ULN if baseline was normal; >3.0 - 10.0×baseline if baseline was abnormal	>10.0×ULN if baseline was normal; >10.0×baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline >ULN. If baseline <LLN, then ranges for "normal baseline" are applied.
CD4 lymphocytes decreased	<LLN - 500/mm^3; <LLN - 0.5×$10e^9$ /L	<500 - 200/mm^3; <0.5 - 0.2×$10e^9$ /L	<200 - 50/mm^3; <0.2 - 0.05×$10e^9$ /L	<50/mm^3; <0.05×$10e^9$ /L	
Cholesterol high	>ULN - 300 mg/dL; >ULN - 7.75 mmol/L	>300 - 400 mg/dL; >7.75 - 10.34 mmol/L	>400 - 500 mg/dL; >10.34 - 12.92 mmol/L	>500 mg/dL; >12.92 mmol/L	
CPK increased	>ULN - 2.5×ULN	>2.5×ULN - 5×ULN	>5×ULN - 10×ULN	>10×ULN	
Creatinine increased	Creatine Kinase >ULN - 1.5×ULN	Creatine Kinase >1.5 - 3.0×baseline; >1.5 - 3.0×ULN	Creatine Kinase >3.0×baseline; >3.0 - 6.0×ULN	Creatine Kinase >6.0×ULN	
Fibrinogen decreased	<1.0 - 0.75×LLN; if abnormal, <25% decrease from baseline	<0.75 - 0.5×LLN; if abnormal, 25% - <50% decrease from baseline	<0.5 - 0.25 x LLN; if abnormal, 50% - <75% decrease from baseline	<0.25×LLN; if abnormal, 75% decrease from baseline; absolute value <50 mg/dL	Ranges defined for "abnormal" are applied only on values <LLN. Grade 0 will be assigned to values >ULN.
GGT increased	>ULN - 2.5×ULN if baseline was normal; 2.0 - 2.5×baseline if baseline was abnormal	>2.5 - 5.0×ULN if baseline was normal; >2.5 - 5.0×baseline if baseline was abnormal	>5.0 - 20.0×ULN if baseline was normal; >5.0 - 20.0×baseline if baseline was abnormal	>20.0×ULN if baseline was normal; >20.0×baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline >ULN. If baseline <LLN, then ranges for "normal baseline" are applied.
Haptoglobin decreased	<LLN	-	-	-	
Hemoglobin increased	Increase in >0 - 2 g/dL; Increase in >0 - 20 g/L	Increase in >2 - 4 g/dL; Increase in >20 - 40 g/L	Increase in >4 g/dL; Increase in >40 g/L	-	The increase indicates the level of increase above normal (above ULN). Applied as, eg grade 1 (g/dL): >ULN – ULN+2 g/dL; Added ranges in SI unit (g/L).

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
INR increased	>1.2 - 1.5; *>1 - 1.5×baseline if on anticoagulation; monitoring only indicated*	>1.5 - 2.5; *>1.5 - 2.5×baseline if on anticoagulation; dose adjustment indicated*	>2.5; *>2.5×baseline if on anticoagulation; bleeding*	-	Concomitant therapy or clinical signs and symptoms are not taken into consideration for grading.
Lipase increased	>ULN - 1.5×ULN	>1.5 - 2.0×ULN; *>2.0 - 5.0×ULN and asymptomatic*	>2.0 - 5.0×ULN with signs or symptoms; *>5.0×ULN and asymptomatic*	>5.0×ULN and with signs or symptoms	"Asymptomatic" ranges are not taken into consideration for grading, ie worst case grading is applied.
Lymphocyte count decreased	<LLN - 800/mm^3; <LLN - 0.8 x 10e^9/L	<800 - 500/mm^3; <0.8 - 0.5×10e^9 /L	<500 - 200/mm^3; <0.5 - 0.2×10e^9 /L	<200/mm^3; <0.2×10e^9 /L	
Lymphocyte count increased	-	>4000/mm^3 - 20,000/mm^3; >4 - 20×10e^9 /L	>20,000/mm^3; >20×10e^9 /L	-	Added ranges in SI unit (×10e^9 /L).
Neutrophil count decreased	<LLN - 1500/mm^3; <LLN - 1.5×10e^9 /L	<1500 - 1000/mm^3; <1.5 - 1.0×10e^9 /L	<1000 - 500/mm^3; <1.0 - 0.5×10e^9 /L	<500/mm^3; <0.5×10e^9 /L	Both Neutrophils and segmented neutrophils are graded using these criteria.
Platelet count decreased	<LLN - 75,000/mm^3; <LLN - 75.0×10e^9 /L	<75,000 - 50,000/mm^3; <75.0 - 50.0×10e^9 /L	<50,000 - 25,000/mm^3; <50.0 - 25.0×10e^9 /L	<25,000/mm^3; <25.0×10e^9 /L	
Serum amylase increased	>ULN - 1.5×ULN	>1.5 - 2.0×ULN; *>2.0 - 5.0×ULN and asymptomatic*	>2.0 - 5.0×ULN *with signs or symptoms;* *>5.0×ULN and asymptomatic*	>5.0×ULN *and with signs or symptoms*	"Asymptomatic" ranges are not taken into consideration for grading, ie worst case grading is applied.
White blood cell decreased	<LLN - 3000/mm^3; <LLN - 3.0×10e^9 /L	<3000 - 2000/mm^3; <3.0 - 2.0×10e^9 /L	<2000 - 1000/mm^3; <2.0 - 1.0×10e^9 /L	<1000/mm^3; <1.0×10e^9 /L	
Metabolism and nutrition disorders					
Acidosis	pH <normal, but >=7.3	-	pH <7.3	*Life-threatening consequences*	pH <normal is implemented as pH <LLN. Clinical signs and symptoms are not taken into consideration for grading.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Alkalosis	pH >normal, but <=7.5	-	pH >7.5	*Life-threatening consequences*	pH >normal is implemented as pH >ULN. Clinical signs and symptoms are not taken into consideration for grading.
Hypercalcemia	Corrected serum calcium of >ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L; Ionized calcium >ULN - 1.5 mmol/L	Corrected serum calcium of >11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L; Ionized calcium >1.5 - 1.6 mmol/L; *symptomatic*	Corrected serum calcium of >12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L; *hospitalization indicated*	Corrected serum calcium of >13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hyperkalemia	Potassium >ULN - 5.5 mmol/L	Potassium >5.5 - 6.0 mmol/L; *intervention initiated*	Potassium >6.0 - 7.0 mmol/L; *hospitalization indicated*	Potassium >7.0 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypermagnesemia	Magnesium >ULN - 3.0 mg/dL; >ULN - 1.23 mmol/L	-	Magnesium >3.0 - 8.0 mg/dL; >1.23 - 3.30 mmol/L	Magnesium >8.0 mg/dL; >3.30 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypernatremia	Sodium >ULN - 150 mmol/L	Sodium >150 - 155 mmol/L; *intervention initiated*	Sodium >155 - 160 mmol/L; *hospitalization indicated*	Sodium >160 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypertriglyceridemia	Triglycerides 150 mg/dL - 300 mg/dL; 1.71 mmol/L - 3.42 mmol/L	Triglycerides >300 mg/dL – 500 mg/dL; >3.42 mmol/L - 5.7 mmol/L	Triglycerides >500 mg/dL – 1000 mg/dL; >5.7 mmol/L - 11.4 mmol/L	Triglycerides >1000 mg/dL; >11.4 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypoalbuminemia	Albumin <LLN - 3 g/dL; <LLN - 30 g/L	Albumin <3 - 2 g/dL; <30 - 20 g/L	Albumin <2 g/dL; <20 g/L	*Life-threatening consequences; urgent intervention indicated*	Clinical signs and symptoms are not taken into consideration for grading.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Hypocalcemia	Corrected serum calcium of <LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L; Ionized calcium <LLN - 1.0 mmol/L	Corrected serum calcium of <8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L; Ionized calcium <1.0 - 0.9 mmol/L; *symptomatic*	Corrected serum calcium of <7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L; Ionized calcium <0.9 - 0.8 mmol/L; *hospitalization indicated*	Corrected serum calcium of <6.0 mg/dL; <1.5 mmol/L; Ionized calcium <0.8 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypoglycemia	Glucose <LLN - 55 mg/dL; <LLN - 3.0 mmol/L	Glucose <55 - 40 mg/dL; <3.0 - 2.2 mmol/L	Glucose <40 - 30 mg/dL; <2.2 - 1.7 mmol/L	Glucose <30 mg/dL; <1.7 mmol/L; *life-threatening consequences; seizures*	Clinical signs and symptoms are not taken into consideration for grading. Urine glucose is not graded.
Hypokalemia	*Potassium <LLN - 3.0 mmol/L*	*Symptomatic with* Potassium <LLN - 3.0 mmol/L; intervention indicated	Potassium <3.0 - 2.5 mmol/L; *hospitalization indicated*	Potassium <2.5 mmol/L; *life-threatening consequences*	"Symptomatic" ranges are applied for grade 2, grade 1 not assigned, ie worst case applied. Clinical signs and symptoms are not taken into consideration for grading of grade 3 and 4.
Hypomagnesemia	Magnesium <LLN - 1.2 mg/dL; <LLN - 0.5 mmol/L	Magnesium <1.2 - 0.9 mg/dL; <0.5 - 0.4 mmol/L	Magnesium <0.9 - 0.7 mg/dL; <0.4 - 0.3 mmol/L	Magnesium <0.7 mg/dL; <0.3 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hyponatremia	Sodium <LLN - 130 mmol/L	*Sodium 125-129 mmol/L and asymptomatic*	*Sodium 125-129 mmol/L symptomatic; 120-124 mmol/L regardless of symptoms* Sodium <130-120 mmol/L	Sodium <120 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading. Worst case ("<130-120 mmol/L" for grade 3 added by Janssen) is applied across grade 2/3 ranges: 120-129 mol/L

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
					assigned to grade 3, grade 2 not used.
Renal and urinary disorders					
Proteinuria	1+ proteinuria; urinary protein ≥ULN - <1.0 g/24 hrs; urinary protein ≥ULN - <1000 mg/day	**Adult:** 2+ and 3+ proteinuria; urinary protein 1.0 - <3.5 g/24 hrs; urinary protein 1000 - <3500 mg/day **Pediatric:** Urine P/C (Protein/Creatinine) ratio 0.5 - 1.9; Urine P/C (Protein/Creatinine) 56.5 – 214.7 g/mol	**Adult:** 4+ proteinuria; urinary protein ≥3.5 g/24 hrs; urinary protein ≥3500 mg/day; **Pediatric:** Urine P/C (Protein/Creatinine) ratio >1.9; Urine P/C (Protein/Creatinine) >214.7 g/mol	-	In case both 24-h urine collection and dipstick are collected, then worst case is taken, as opposed to having 24-h urine collection take precedence over dipstick. Added ranges in SI unit for urinary protein (mg/day) and for urine P/C (g/mol). Pediatric grading is applied to age range [0-18]. Adult grading is applied for ages [>18].

* Grade 0 is assigned to a lab assessment when the lab test is described in the table, but the lab value is not assigned a grade 1 or higher.

6.10. Appendix 10: Statistical Hypothesis

Hypothesis	Control
Primary Endpoints	
H1. Guselkumab 100 mg q4w SC is superior to placebo as assessed by the proportion of subjects achieving an ACR 20 response at Week 24 ***(primary hypothesis)***	Controlled as in Figure 2
H2. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by the proportion of subjects achieving an ACR 20 response at Week 24	Controlled as in Figure 2
Major Secondary Endpoints	
H3. Guselkumab 100 mg q4w SC is superior to placebo as assessed by proportion of subjects who achieved a psoriasis IGA response at Week 24 among the subjects with ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline	Controlled as in Figure 2
H4. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by proportion of subjects who achieved a psoriasis IGA response at Week 24 among the subjects with ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline	Controlled as in Figure 2
H5. Guselkumab 100 mg q4w SC is superior to placebo as assessed by proportion of subjects who achieved PASI 90 response at Week 24	Controlled as in Figure 2
H6. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by proportion of subjects who achieved PASI 90 response at Week 24	Controlled as in Figure 2
H7. Guselkumab 100 mg q4w SC is superior to placebo as assessed by change from baseline in HAQ-DI score at Week 24	Controlled as in Figure 2
H8. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by change from baseline in HAQ-DI score at Week 24	Controlled as in Figure 2

H9. Guselkumab 100 mg q4w SC is superior to placebo as assessed by change from baseline in SF-36 PCS at Week 24	Controlled as in Figure 2
H10. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by change from baseline in SF-36 PCS at Week 24	Controlled as in Figure 2
H11. Guselkumab 100 mg q4w SC is superior to placebo as assessed by change from baseline in Facit-F at Week 24	Controlled as in Figure 2
H12. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by change from baseline in Facit-F at Week 24	Controlled as in Figure 2
H13. Guselkumab 100 mg q4w SC is superior to placebo as assessed by proportion of subjects who achieved MDA at Week 24	Controlled as in Figure 2
H14. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by proportion of subjects who achieved MDA at Week 24	Controlled as in Figure 2
H15. Guselkumab 100 mg q4w SC is superior to placebo as assessed by proportion of subjects who achieved an ACR 50 response at Week 24	Weakly-controlled
H16. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by proportion of subjects who achieved an ACR 50 response at Week 24;	Weakly-controlled
H17. Guselkumab 100 mg q4w SC is superior to placebo as assessed by proportion of subjects achieved an ACR 20 response at Week 16	Weakly-controlled
H18. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by proportion of subjects achieved an ACR 20 response at Week 16	Weakly-controlled
H19. Guselkumab 100 mg q4w SC is superior to placebo as assessed by proportion of subjects who achieve an ACR 70 response at Week 24	Weakly-controlled
H20. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by	Weakly-controlled

proportion of subjects who achieve an ACR 70 response at Week 24	
H21. Guselkumab 100 mg q4w SC is superior to placebo as assessed by proportion of subjects who achieve an ACR 50 response at Week 16	Weakly-controlled
H22. Guselkumab 100 mg at Week 0, Week 4, then q8w SC is superior to placebo as assessed by proportion of subjects who achieve an ACR 50 response at Week 16	Weakly-controlled

For hypothesis testing order and multiplicity adjustment, refer to Section 5.3.3.2.

6.11. Appendix 11: Summary Rules Applied in Definitions of Endpoints

1. Joint Evaluability Rules for Sign and Symptom Data

For participants having a joint injection(s)/surgical joint procedure(s) prior to the date of study entry (eg, randomization) or during the study, the affected joint(s) will be valued according to the following rules:

- For participants having a joint injection and/or surgical joint procedure prior to the date of randomization, the affected joints will be analyzed according to the impact of the joint injection and/or surgical joint procedure on the evaluability of the involved joints.
- If a joint is considered un-evaluable at baseline due to certain procedure/injection performed prior to the date of randomization, the joint will be considered un-evaluable throughout the study.
- For participants undergoing joint procedures for the treatment of PsA during the study, the affected joints will be considered as swollen and tender from the date of procedure onwards.
- For participants undergoing joint procedures during the study for the treatment of non-PsA disease indication, the affected joints will be analyzed according to the impact of the surgical joint procedure on the evaluability of the involved joints.
- For participants undergoing joint injections for PsA during the study, the affected joints will be considered as swollen and tender from the date of injection for the next 90 days.
- For participants undergoing joint injections for non-PsA related reasons during the study, the affected joints will be considered as non-evaluable from the date of injection for the next 90 days.

2. Joint Count Adjustment Rule

For participants who have an incomplete set of evaluable joints the joint count/score will be adjusted to the total number joints of interest (eg, 68 joints for tenderness and 66 joints for swelling) by dividing the number of affected joints by the number of evaluable joints and multiplying by the total number joints of interest.

3. LLOQ Rule

Any value <LLOQ is considered equal to half of the value of LLOQ for numerical calculations.

4. Joint Evaluability Rules

A SCREENING x-ray is used to confirm sacroiliitis. (Only applies to patients that present with peripheral arthritis as their primary arthritic presentation of PsA)

Upon confirmation the subjects are asked to complete BASDAI questionnaire at Week 0, 8, 16, 24, 52, 76 and 100.

If sacroiliitis is not confirmed, the subjects will NOT complete the BASDAI questionnaires. However, subject WITHOUT confirmed sacroiliitis will still participate in all other study assessments.

6.12. Appendix 12: Summary of Analyses Based on Multiple Imputation

Summary of Multiple Imputation Method

Endpoints (Population[a]) Estimand[b]	MI specification	Analysis method/Summary statistics
ACR20 at Week 24 ACR20 at Week 16 ACR50 at Week16 and Week24 ACR70 at Week24 HAQ-DI Score (FAS) Adjusted Composite Estimand (FAS-UKR) Treatment Policy Estimand	Multiple imputation with FCS regression of component scores	MIdata1 (N=200, Seed[d]=18496) • Imputation variables: 7 ACR components from Week 0-Week24 • Ancillary variables: Intervention group, baseline non-biologic DMARD use (yes/no)
PASI 90 response at Week 24 among subjects with ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline (FAS) Adjusted Composite Estimand (FAS-UKR) Treatment Policy Estimand	Multiple imputation with FCS regression of component scores	MIdata2 (N=200, Seed[d]=6509723) Note: The MI is done only for the subset of subjects who have ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline, and not all subjects in the study population. • Imputation variables: PASI score from Weeks 0-24 • Ancillary variables [c]: Intervention group, baseline non-biologic DMARD use (yes/no), 7 ACR components from Weeks 0–24
Achieving MDA At week24 (FAS) Adjusted Composite Estimand (FAS-UKR) Treatment Policy Estimand	Multiple imputation with FCS regression of component scores	MIdata3 (N=200, Seed[d]=6509723) • Imputation variables: 7 MDA components from Week 0-24 • Ancillary variables: Intervention group, baseline non-biologic DMARD use (yes/no)
Change from baseline through Week 24 in: • PCS score • Facit-F score (FAS) Adjusted Composite Estimand (FAS-UKR) Treatment Policy Estimand	Multiple imputation with FCS regression of PC	MIdata4 (N=200, Seed[d] =890473) • Imputation variables: PCS scores and Facit-F scores from Weeks 0–24 • Ancillary variables [c]: Intervention group, baseline non-biologic DMARD use (yes/no), 7 ACR components from Weeks 0-24
IGA change from baseline through Week 24 among subjects with ≥3% BSA psoriatic involvement and an IGA score of ≥2 at baseline (FAS) Adjusted Composite Estimand (FAS-UKR) Treatment Policy Estimand	Multiple imputation with FCS regression of IGA scores	MIdata5 (N=200, Seed[d] = 413249) Note: The MI is done only for the subset of subjects who have ≥3% BSA psoriatic involvement and an IGA score of ≥2 (mild) at baseline, and not all subjects in the study population. • Imputation variables: IGA scores from Weeks 0–24 Ancillary variables [c]: Intervention group, baseline non-biologic DMARD use (yes/no) 7 ACR components from Weeks 0-24

[a] The population defines which subset of participants the imputation will be performed for.

[b] The handling of ICEs associated with the estimand listed will be applied to the imputation variables and ancillary variables (if post-baseline) prior to imputation.
[c] An ancillary variable may be removed if its correlation with an indicator variable that determines missing-ness of the variable to be imputed is low or, there are too many missing values for the ancillary variable within the subgroup of incomplete cases for the variable to be imputed. All the 7 ACR components (including all measurements from baseline through Week 24) are in the list of the ancillary variables since they may be related to the mechanism leading to missing data.
[d] The starting seed for FCS regression MI is used to generate a series of imputation seeds using the algorithm: INT((2**31-2)*RANUNI(starting seed)), where each imputation seed will be used for a single imputation. To account for the possibility that some imputations may fail to complete due to out-of-range issues, 200+ initial imputation seeds will be prepared, and the first 200 successful imputations will be used for analysis.

6.13. Appendix 13: Description of Statistical Models

MMRM Model

To account for the missing data for continuous endpoints, an MMRM model will be used on the change from baseline, under the assumption of MAR, to test the difference between a guselkumab group and the placebo group. The model will include treatment group, the interaction terms of visit with treatment group, baseline non-biologic DMARD use (yes/no), and baseline score as explanatory variables. An unstructured covariance matrix for repeated measure within a subject will be used. The F-test will use Kenward-Roger's approximating for degree of freedom. In case of lack of convergence, empirical structured covariances will be used in the following order until convergence is reached: 1) Toeplitz 2) first order Autoregressive Moving Average. For analyses through Week 24 the model will include data from all 3 treatment groups through Week 24. The treatment difference between a guselkumab group and the placebo group will be estimated by the difference in the least squares means (LSmeans). The 95% CI for the differences in LSmeans and p-values will be calculated based on the MMRM.

After Week 24, the MMRM model may still be used to generate LSmeans for each treatment group. However, LSmeans difference between treatments and associated p-values will no longer be calculated.

ANCOVA Model

The ANCOVA model will be used on the change from baseline, under the assumption of MCAR, to test the difference between a guselkumab group and the placebo group. The model will include treatment group, baseline non-biologic DMARD use (yes/no), and baseline score as explanatory variables. For analyses through Week 24 the model will include data from all 3 treatment groups at Week 24. The treatment difference between a guselkumab group and the placebo group will be estimated by the difference in the LSmeans. The 95% CI for the differences in LSmeans and p-values will be calculated based on ANCOVA.

After Week 24, the ANCOVA model may still be used to generate LSmeans for each treatment group. However, LSmeans difference between treatments and associated p-values will no longer be calculated.

GLMM

To account for the missing data for binary endpoints, a GLMM will be used on the response status, under the assumption of MAR, to test the difference between a guselkumab group and the placebo group. The model will include treatment group, the interaction terms of visit with treatment group, and baseline non-biologic DMARD use (yes/no) as explanatory variables. An unstructured covariance matrix for repeated measure within a subject will be used. The logit link will serve as the link function. In case of lack of convergence, empirical structured covariances will be used in the following order until convergence is reached: 1) Toeplitz 2) first order Autoregressive Moving Average. For analyses through Week 24 the model will include data from all 3 treatment groups through Week 24. The proportion difference, its 95% CI, and p-value will be calculated based on the GLMM.

After Week 24, the GLMM may still be used to generate proportion of response for each treatment group. However, the proportion difference between treatments and associated p-values will no longer be calculated.

6.14. Appendix 14: Summary of ICEs and ICE Strategy by Estimand

	ICE Category	ICE Strategy by Estimand	
		Adjusted Composite	**Treatment Policy**
1	Discontinued study intervention injections due to any reason except Natural Disaster or Major Disruption	• Composite strategy • Consider all subsequent data TF • Priority = high	• Treatment Policy strategy • No action taken, use all observed data
2	Initiated or increased the dose of non-biologic DMARDs (MTX, SSZ, HCQ, LEF) or oral corticosteroids over baseline for PsA	• Composite strategy • Consider all subsequent data TF • Priority = high	
3	Initiated protocol prohibited medications/ therapies for PsA	• Composite strategy • Consider all subsequent data TF • Priority = high	
4	Discontinued study intervention injections due to due to Natural Disaster or Major Disruption	• Hypothetical strategy • Do not use all subsequent data (assuming MAR) • Priority = low	
5	Severe treatment non-compliance due to Natural Disaster or Major Disruption. Defined as: for a given visit, when the total number of doses of active (ie, guselkumab) study intervention is missed due to Natural Disaster or Major Disruption exceeds 30% of the total protocol defined active doses from Week 0 up to and including that visit. For Weeks 0 to 20, this amounts to ≥2 dose missed	• Hypothetical strategy • Do not use data at subsequent visit (assuming MAR) • Priority = low	

Note: For continuous endpoints, participants considered TF should be set as change=0; for binary endpoints, participants considered TF should be set as NR.
Note: High priority ICE strategies will supersede low priority. ICE strategies should both be applicable to the same visit.

Definitions:
Natural Disaster: site closure, site access restrictions, or lockdowns caused by COVID-19.
Major Disruption: the disruption involving Ukraine and neighboring countries/territories beginning 24 February 2022.

7. REFERENCES

Cella D, Yount S, Sorensen M, et al. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. J Rheum.2005;32(5):811-819.

Chandran V, Bhella S, Schentag C, Gladman D. Functional assessment of chronic illness therapy-fatigue scale is valid in patients with psoriatic arthritis. Ann Rheum Dis. 2007;66:936-939.

Coates LC, Helliwell PS. Validation of minimal disease activity criteria for psoriatic arthritis using interventional data. Arthritis Care Res. 2010;62(7):965-969.

Di Carlo, M, Becciolini A, Lato V, et al. The 12-item Psoriatic Arthritis Impact of Disease Questionnaire: Construct Validity, Reliability, and Interpretability in a Clinical Setting. J Rheumatol. 2017;44(3):279-285.

Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum. 1993;36(6):729-740.

Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum. 1995;38(6):727-735.

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

Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcomes in arthritis. Arthritis Rheum. 1980;23(2):137-145.

Gladman DD, Inman RD, Cook RJ, et al. International spondyloarthritis interobserver reliability exercise--the INSPIRE study: II. Assessment of peripheral joints, enthesitis, and dactylitis. J Rheumatol. 2007;34(8):1740-1745.

Gladman DD, Ziouzina O, Thavaneswaran A, Chandran V. Dactylitis in psoriatic arthritis: prevalence and response to therapy in the biologic era. J Rheumatol. 2013;40(8):1357-1359.

Healy PJ, Helliwell PS. Measuring clinical enthesitis in psoriatic arthritis: assessment of existing measures and development of an instrument specific to psoriatic arthritis. Arthritis Rheum. 2008;59(5):686-691.

Helliwell PS, Kavanaugh A. Comparison of composite measures of disease activity in psoriatic arthritis using data from an interventional study with golimumab. Arthritis Care Res. 2014;66(5):749-756.

Kim Y, Won S. Adjusted proportion difference and confidence interval in stratified randomized trials. PharmaSUG 2013 – Paper SP04.

Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum. 2000;43(7):1478–1487.

Mundt JC, Greist JH, Jefferson JW, et al. Prediction of suicidal behavior in clinical research by lifetime suicidal ideation and behavior ascertained by the electronic Columbia-Suicide Severity Rating Scale. J Clin Psychiatry. 2013;74(9):887-893.

Pham T, Guillemin F, Claudepierre P, et al. TNFα antagonist therapy in ankylosing spondylitis and psoriatic arthritis: recommendations of the French Society for Rheumatology. Joint Bone Spine. 2006;73:547–553.

Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011;168(12):1266-1277.

Ratitch B, Lipkovich I, O'Kelly M. Combining analysis results from multiply imputed categorical data. PharmaSUG 2013-Paper SP03.

Rubin, DB. Multiple Imputations for Nonresponse in Surveys. New York: John Wiley & Sons. 1987.

Saris-Bagmana RN, Dewey CJ, Chisholm GB, et al. QualityMetric Health OutcomesTM Scoring Software 4.5 User's Guide. QualityMetric Health Outcomes Solutions. 2011.

Smolen JS, Beaulieu A, Rubbert-Roth A, et al, for the OPTION Investigators. Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. Lancet. 2008;371(9617):987-997.

Ware JE Jr, Sherbourne CD. The MOS 36 item short form health survey (SF 36), I: conceptual framework and item selection. Med Care. 1992;30(6):473-483.

Yount S, Sorensen MV, Cella D, et al. Adalimumab plus methotrexate or standard therapy is more effective than methotrexate or standard therapies alone in the treatment of fatigue in patients with active, inadequately treated rheumatoid arthritis. Clin Exp Rheumatol. 2007;25(6):838-46.