Statistical Analysis Plan J1A-MC-KDAF (3)

A Phase 2b, Double-blinded, Placebo-Controlled Study to Evaluate Peresolimab in Adult Participants with Moderately-to-Severely Active Rheumatoid Arthritis

NCT05516758

Approval Date: 01-Dec-2024

Title Page

Protocol Title: A Phase 2b, Double-blinded, Placebo-Controlled Study to Evaluate Peresolimab in Adult Participants with Moderately-to-Severely Active Rheumatoid Arthritis

Protocol Number: J1A-MC-KDAF

Compound Number: Peresolimab (LY3462817)

Short Title: A Phase 2b Study to Evaluate Peresolimab in Adult Participants with Moderately-to-Severely Active Rheumatoid Arthritis

Sponsor Name: Eli Lilly and Company

Legal Registered Address: Indianapolis, Indiana, USA 46285

Regulatory Agency Identifier Number(s)

Registry	ID
IND	150710
EU Trial Number	2022-501425-20-00

Document ID: VV-CLIN-131347

Table of Contents

Table of Contents

Tables **Page**

Table of Contents

Figure **Page**

Version History

Statistical Analysis Plan (SAP) Version 1 for Study J1A-MC-KDAF (KDAF) is based on the protocol dated 18 Jul 2022 and approved prior to unblinding.

Statistical Analysis Plan (SAP) Version 2 based on the protocol dated 31 May 2023 was approved prior to the primary outcome database lock and includes the following changes. Minor corrections/additions may not be included.

Section	Description of Change	Rationale
Section 1.2	• Updated participant population	• As per the protocol amendment (e)
Section 4.1	• Added definition for safety analyses.	• Safety analyses will be summarized by period, which requires baseline for each period.
	• Added statement "If ET visit does not occur, last injection date will be considered as the end of the treatment period"	• Clarification
	• Added definition for analysis visit and data to be used for the analysis for eCOA assessments	• To allow to use eCOA assessments collected within the visit window.
	• Added "up to Week 24" and removed "within the same treatment period" for LOCF	• To facilitate analysis in combined periods, Period 2 and Period 3A.
Section 4.4.2.1	• Added methods handling missing data	• Clarification
Section 4.4.3.2	• Added estimand strategy and handling missing data	• Clarification
Section 4.5	• Added estimand strategy and handling missing data	• Clarification
Section 4.6	• Added periods and population for safety summary.	• Safety analyses will be summarized by period.
	• Added statement "For Period 3B, the treatment groups will be combined and not be separated by populations"	• Clarification
Section 4.6.2	• Added definitions of baseline for Period 2 and other periods.	• Baseline definition will be different between Period 2 and other periods.
	• Added additional condition for TEAE. AEs occurring on the	• AE data do not include time of onset.

	date of first injection are considered as TEAE. AEs occurring on date of other injections are considered as post-injection except for the first injection date.	
	• Removed paragraph for common TEAEs with ≥1%.	• To avoid redundancy due to the sample size per treatment group.
	• Added definition of AESI.	• Clarification
Section 4.6.2.2	• Removed "Event maps to PT of Infusion related reaction"	• Not relevant to the trial.
Section 4.6.2.4	• Removed analyses for infections categorized Gram-positive/negative bacterial infection, mixed aerobic-anaerobic bacterial infection, fungal infections, viral infections, other infections, and opportunistic infections as assessed by the investigator.	• Not relevant to the trial.
Section 4.6.2.4	• Added detailed description for opportunistic infections	• Clarification
	• Added analyses plan for herpes simplex.	• Clinical request
Section 4.6.2.5	• Added SMQ for malignancies	• Clarification
Section 4.6.2.6	• Added SMQ for depression	• Clarification
Section 4.6.3	• Added categorization for shifts of hepatic measures, hematology, lipid effects, and renal function.	• Clarification
Section 4.6.4.4	• Added definitions for immunogenicity analysis	• Clarification
	• Removed time to event analysis	• Not informative analysis given the sample size
Section 4.7.4	• Added analysis plan for Japanese subgroup analysis	• To prepare Japanese submission.
Appendix 4	• Added definition of handling missing data for preexisting conditions and medical history	• Clarification

Appendix 5	• Added definition of handling missing data for prior and concomitant medication	• Clarification
Appendix 10	• Added list of PTs for herpes zoster, herpes simplex, and hepatitis B virus	• To provide a full list of PTs.
Appendix 11	• Added definition of grade for hematologic events, lipid effects, and renal function	• To provide categorization for shift tables.

Statistical Analysis Plan (SAP) Version 3 based on the protocol dated 31 May 2023 was approved prior to the final database lock and includes the following changes. Minor corrections/additions may not be included.

Section	**Description of Change**	**Rationale**
Section 1.2 & 4.1	• Replaced "early termination" to "early discontinuation", and "ET" to "ED"	• To use the consistent terminology as per study protocol.
Section 3 – Table KDAF.3.1	• Removed mITT population, added safety population for exposure table and safety population who received at least 1 dose of peresolimab to the Treatment Period (combined from Periods 2 and 3; Weeks 0 to 60), and updated the treatment groups	• Selected safety analyses are to be performed using this combined period.
Section 4.1	• Added the reason why analysis visit numbers for eCOA assessments are to be re-assigned.	• Clarification
Section 4.4.1	• Clarified that permanently discontinuing study intervention is one of the intercurrent events related study invention.	• Clarification
Section 4.5 – Table KDAF.4.1	• Added references/citations to some efficacy and patient reported outcome measures	• Clarification

Section 4.5 – Table KDAF.4.1	• Replaced all references of "PsA" to "RA"	• Correction
Section 4.5 – KDAF.4.2	• Updated the MRI analysis approach from "LOCF" to "BOCF"	• Correction
Section 4.6.2	• Added the instructions on how to identify AE leading to temporary interruption of study drug	• Clarification
Section 4.6.2	• Clarified the analysis of temporary interruption of study drug is based on interruptions due to AEs only and added the duration calculation of study drug interruption	• Clarification
Section 4.6.2.1	• Removed the listing of AEs occurred >50 days after study treatment discontinuation	• These AEs are already included in the listing of AEs. Separate listing is not needed.
Section 4.6.2.2	• Clarified the frequency of treatment-emergent hypersensitivities is to be summarized by treatment arm, SMQ and PT	• Clarification
Section 4.6.2.2	• Provided clarification on how the exposure-adjusted incidence rates are calculated	• Clarification
Section 4.6.2.3	• Added the missing categories for erythema	• Correction
Section 4.6.2.4	• Removed the unnecessary analyses, such as infectious AE for patients who are treated with antibiotics, infectious AE by infection diagnosis, treatment with antiviral medication for herpes zoster and herpes simples, and detailed listing of infectious AEs.	• Not relevant to study
Section 4.6.2.5	• Added a new SMQ code in addition to the current	• Per GPS request

	Malignancy SMQ code; added the list of PTs categorized as NMSC vs. non-NMSC; and added the time to onset analysis for malignancy	
Section 4.6.4.2	• Removed the analyses on Ig levels and cytokines with ADA	• Not relevant
Section 4.6.2.5 & 4.6.2.7	• Removed the graphs of exposure-adjusted incidence rates for malignancy and depression	• Correction
Section 4.6.3	• Updated the AST, AST, ALP and Total bilirubin threshold for the shift from baseline analysis	• Correction
Section 4.6.3	• Specified the shift from baseline analyses of hematologic changes, lipid effects and renal function are to be performed using maximum baseline and postbaseline results	• Clarification
Section 4.6.4.2	• Removed the analyses which are not needed.	• Correction
Section 4.7.1	• Clarified the subgroups for the efficacy subgroup analyses	• Clarification
Section 6.2	• Added the definition of overall compliant	• Clarification
Section 6.8	• Removed Italy from the list of geographic region since no subjects from Italy were enrolled.	• Correction

Table KDAF.1.1. SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	07 Apr 2023	Not Applicable	Original version
2	29 Nov 2023	See above	Version 2
3	See date on Page 1		

Abbreviation: SAP = statistical analysis plan.

1. Introduction

Study J1A-MC-KDAF (KDAF) is a Phase 2b, double-blind, placebo-controlled study to evaluate peresolimab in adult participants with moderately-to-severely active rheumatoid arthritis (RA).

There are no changes to the analyses described in the protocol.

1.1. Objectives, Endpoints, and Estimands

Objectives	Endpoints
Primary	
To evaluate the efficacy of peresolimab compared to placebo in achieving ACR20	• Proportion of participants achieving ACR20 at Week 12
Secondary	
To evaluate the effect of peresolimab compared to placebo for measures of disease activity	• Proportion of participants achieving ACR50 or ACR70 at Week 12 • Proportion of participants achieving LDA or remission at Week 12 for o DAS28-hsCRP o DAS28-ESR o SDAI, and o CDAI. • Change from baseline at Week 12 for mean o DAS28-hsCRP o SDAI, and o CDAI. • Change from baseline at Week 12 for ACR core set values o 68 tender joint count o 66 swollen joint count, and o Physician's Global Assessment of Disease Activity (VAS).
To evaluate the effect of peresolimab compared to placebo for patient-reported outcome measures	• Change from baseline at Week 12 for patient-reported ACR core set values o Patient's Global Assessment of Disease Activity (VAS) o Patient's Assessment of Arthritis Pain (VAS), and o Patient's Assessment of Physical Function using HAQ-DI. • Change from baseline at Week 12 for o the duration and severity of morning joint stiffness and o Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) scores. • Change from baseline at Week 12 for o SF-36 domains o SF-36 Physical Component Summary, and o SF-36 Mental Component Summary.

Objectives	Endpoints
To evaluate the effect of peresolimab for measures of disease activity	• Proportion of participants achieving LDA or remission at Weeks 24 and 60 for o DAS28-hsCRP o DAS28-ESR o SDAI, and o CDAI. • Proportion of participants achieving ACR20, ACR50, or ACR70 at Weeks 24 and 60. • Change from baseline at Weeks 24 and 60 for ACR core set values o 68 tender joint count o 66 swollen joint count, and o Physician's Global Assessment of Disease Activity (VAS). • Change from baseline at Weeks 24 and 60 for mean o DAS28-hsCRP o SDAI, and o CDAI.
To characterize the pharmacokinetics of peresolimab	• Observed trough drug concentration at Week 12
Exploratory	
CCI	

Objectives	Endpoints
CCI	

Abbreviations: ACR = American College of Rheumatology; CDAI = Clinical Disease Activity Index; CCI DAS28-hsCRP = Disease Activity Score – high-sensitivity C-reactive protein; CCI HAQ-DI = Health Assessment Questionnaire - Disability Index; LDA = low disease activity; CCI VAS = visual analog scale.

Primary estimand

The primary clinical question of interest is:

> What is the difference between peresolimab and placebo in the target patient population, in achieving a successful response at Week 12 without use of rescue or prohibited medications or discontinuing the study intervention?

The estimand is described by the following attributes:

- Population: Participants with moderately-to-severely active RA
- Endpoint: American College of Rheumatology (ACR) 20 at Week 12
- How to account for intercurrent events (ICEs): ICEs related to study intervention include the use of rescue or prohibited medication and early discontinuation (ED) from the study or study intervention. A composite strategy will be used for these types of ICEs. Specifically, participants with ICEs related to study intervention will be considered as treatment failure (having no change from baseline), that is, a nonresponder, after the first occurrence of these ICEs.
- Population-level summary: Difference in proportion of participants achieving response at Week 12 between each dosing regimen of peresolimab and placebo
- Rationale for estimand: The primary estimand strategy assumes that:

- If a participant used any rescue or prohibited medication, the participant was not receiving sufficient benefits from study intervention.
- If a participant early discontinued the study or study intervention, the participant experienced a burden of study intervention that outweighed its benefits.

Secondary estimand for categorical endpoints at Week 12

For secondary objectives that have categorical endpoints analyzed at Week 12, the clinical question of interest is:

> What is the difference between peresolimab and placebo in the target patient population, in achieving a successful response at Week 12 without use of rescue or prohibited medications or discontinuing the study intervention?

The estimand is described by the following attributes:

- Population: Participants with moderately-to-severely active RA
- Endpoint(s): ACR50, ACR70, Disease Activity Score – high-sensitivity C-reactive protein (DAS28-hsCRP) low disease activity (LDA), Disease Activity Score – erythrocyte sedimentation rate (DAS28-ESR) LDA, Simplified Disease Activity Index (SDAI) LDA, Clinical Disease Activity Index (CDAI) LDA, DAS28-hsCRP remission, DAS28-ESR remission, SDAI remission, and CDAI remission at Week 12
- How to account for ICEs: ICEs will be accounted using the same estimand strategy as for the primary estimand.
- Population-level summary: Difference in proportion of participants achieving response at Week 12 between each dosing regimen of peresolimab and placebo
- Rationale for estimand: Refer to the rationale for the primary estimand.

Secondary estimand for continuous endpoints at Week 12

For secondary objectives that have continuous endpoints analyzed at Week 12, the clinical question of interest is:

> What is the difference between peresolimab and placebo in the target patient population, in average change from baseline to Week 12 if all participants continued with treatment for 12 weeks without use of rescue or prohibited medications?

The estimand is described by the following attributes:

- Population: Participants with moderately-to-severely active RA
- Endpoint(s): Change from baseline for DAS28-hsCRP, SDAI, CDAI, 68 tender joint count (TJC68), 66 swollen joint count (SJC66), Physician's Global Assessment of Disease Activity (visual analog scale [VAS]), Patient's Global Assessment of Disease Activity (VAS), Patient's Assessment of Arthritis Pain (VAS), patient's assessment of physical function using the Health Assessment Questionnaire - Disability Index (HAQ-DI), the duration and severity of morning joint stiffness, and Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) scores at Week 12

- How to account for ICEs: ICEs related to study intervention include the use of rescue or prohibited medication and ED from the study or study intervention. A hypothetical strategy will be used for these types of ICEs. Specifically, data collected after the first occurrence of ICE related to study intervention will be excluded from an analysis, and what the treatment effect would have been if rescued or prohibited medication(s) was not used, and all patients adhered to the treatment for 12 weeks will be estimated.
- Population-level summary: Difference in mean change from baseline to Week 12 between intervention conditions
- Rationale for estimand: The data collected after the treatment discontinuation, or the use of rescue or prohibited medication will not present the true efficacy effects.

Secondary estimand for categorical endpoints at Week 24 and Week 60

For secondary objectives that have categorical endpoints analyzed at Week 24 (or Week 60), the clinical question of interest is:

> What is the proportion of achieving a successful response at Week 24 (or Week 60) in the target patient population without use of rescue or prohibited medications or discontinuing the study intervention?

The estimand is described by the following attributes:

- Population: Participants with moderately-to-severely active RA
- Endpoint(s): ACR20, ACR50, ACR70, DAS28-hsCRP LDA, DAS28-ESR LDA, SDAI LDA, CDAI LDA, DAS28-hsCRP remission, DAS28-ESR remission, SDAI remission, and CDAI remission at Week 24 (or Week 60)
- How to account for ICEs: ICEs will be accounted using the same estimand strategy as for the primary estimand.
- Population-level summary: Proportion of participants achieving response at Week 24 (or Week 60) in each dosing regimen of peresolimab
- Rationale for estimand: Refer to the rationale for the primary estimand.

Secondary estimand for continuous endpoints at Week 24 and Week 60

For secondary objectives that have continuous endpoints analyzed at Week 12, the clinical question of interest is:

> What is the average change from baseline to Week 24 (or Week 60) in the target patient population if all participants continued with treatment for 12 weeks without use of rescue or prohibited medications?

The estimand is described by the following attributes:

- Population: Participants with moderately-to-severely active RA
- Endpoint(s): Change from baseline for DAS28-hsCRP, SDAI, CDAI, TJC68, SJC66, and Physician's Global Assessment of Disease Activity (VAS) at Week 24 (or Week 60)

- How to account for ICEs: ICEs will be accounted using the same estimand strategy as for the secondary estimand for continuous endpoint at Week 12. It will be assumed that the same level of efficacy prior to the first ICE will continue, that is, last observation carried forward (LOCF) will be used for imputation.
- Population-level summary: Mean change from baseline to Week 24 in each intervention conditions
- Rationale for estimand: Refer to the rationale for the secondary estimand for continuous endpoints at Week 12.

1.2. Study Design

Study KDAF is a Phase 2b, randomized, double-blind, placebo-controlled study to determine the appropriate dose and dosing frequency of peresolimab in adult participants with moderately-to-severely active RA.

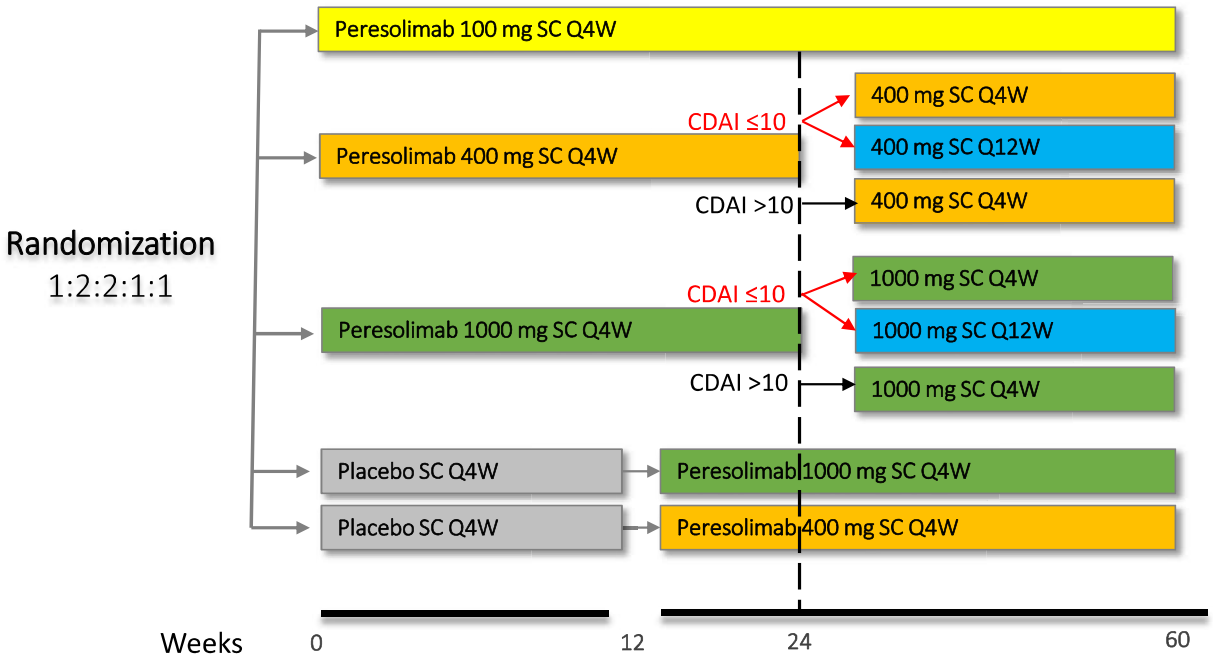
This study includes a

- Period 1: screening period occurring approximately 1 to 42 days prior to Period 2
- Period 2: 12-week double-blind, placebo-controlled treatment period
- Period 3: 48-week double-blind treatment period, and
- Period 4: posttreatment follow-up period occurring 4 and 12 weeks from last treatment period visit (Week 60) or Early Discontinuation Visit.

See the Schedule of Activities (SoA) in the protocol for visit schedule, procedural and assessment details.

In the remainder of this statistical analysis plan (SAP), the first 12 weeks in Period 3 (that is, Weeks 12 to 24) will be referred as Period 3A and the last 36 weeks in Period 3 (that is, Weeks 24 to 60) as Period 3B for an analysis purpose. Combined Periods 2 and 3 will be referred as the Treatment Period (Weeks 0 to 60).

Figure KDAF.1.1 describes the study design for the treatment periods (Periods 2 and 3).



Abbreviations: CDAI = Clinical Disease Activity Index; SC = subcutaneous injection; Q4W = every 4 weeks; Q12W = every 12 weeks.
Note: At Week 12, participants receiving placebo Q4W will begin receiving either 400 mg or 1000 mg peresolimab Q4W. At Week 24, participants receiving 400 mg or 1000 mg peresolimab Q4W may be switched to Q12W dosing following assessment of CDAI scores.

Figure KDAF.1.1 Study design.

Participant population

CCI of the participants randomly assigned to study intervention will have a history of CCI

of participants will have a history of CCI .

Treatment period design

CCI

- 100 mg of peresolimab every 4 weeks (Q4W)
- 400 mg of peresolimab Q4W
- 1000 mg of peresolimab Q4W
- placebo Q4W up to Week 12 followed by 400 mg of peresolimab Q4W, and
- placebo Q4W up to Week 12 followed by 1000 mg of peresolimab Q4W.

Participants randomly assigned to placebo Q4W at Week 0

At Week 12, participants receiving placebo Q4W will begin receiving either 400 mg or 1000 mg peresolimab Q4W based on randomization at Week 0. Treatment assignment will be blinded, and the participants will continue their assigned peresolimab dose and dosing frequency until the end of the study.

Participants randomly assigned to 100 mg peresolimab at Week 0

Participants receiving 100 mg peresolimab will continue their assigned dose and dosing frequency throughout the duration of the treatment period. No changes will occur.

Participants randomly assigned to 400 mg or 1000 mg peresolimab at Week 0

Participants receiving 400 mg or 1000 mg peresolimab will continue their assigned dose and dosing frequency until Week 24. At Week 24, participants may be switched to every 12 weeks (Q12W) dosing following assessment of CDAI scores, unless they received rescue medication prior to Week 24.

Abbreviations: CDAI = Clinical Disease Activity Index; Q4W = every 4 weeks; Q12W = every 12 weeks.

CCI

- At Week 24 using the Visit 9 (Week 24) assessments, a participant does not achieve CDAI LDA
    - These participants will receive permitted rescue therapy.
- For participants that achieved CDAI LDA at Week 24, those who do not maintain CDAI LDA for 2 consecutive visits after Week 24
    - These participants will receive peresolimab Q4W and permitted rescue therapy.
        - If a participant is currently receiving peresolimab Q4W, they will continue receiving peresolimab Q4W.

- If a participant is currently receiving peresolimab Q12W, they will switch to peresolimab Q4W.

CCI [redacted]

Escape criteria after Week 24

A participant not achieving ≥20% improvement in both TJC and SJC at 2 consecutive visits after Week 24, while receiving rescue therapy, should be permanently discontinued from study intervention. Participants may continue in the study and receive additional standard of care therapies at the investigator's discretion.

Participation in an open-label extension study

If an open-label extension study is enrolling at the time the participant completes Visit 18 (Week 60), they will be given the opportunity to enroll in that study.

1.2.1. Method of Assignment to Treatment

CCI [redacted]

- 100 mg of peresolimab Q4W
- 400 mg of peresolimab Q4W
- 1000 mg of peresolimab Q4W
- placebo Q4W up to Week 12 followed by 400 mg of peresolimab Q4W, and
- placebo Q4W up to Week 12 followed by 1000 mg of peresolimab Q4W.

CCI [redacted]

Participants who are randomized to either peresolimab 400 mg Q4W or peresolimab 1000 mg Q4W at Week 0 and meet CDAI ≤10 at Week 24 will be randomly assigned to either Q4W or Q12W in the same dose level with a 1:1 ratio. Participants will be stratified by

- CDAI ≤2.8 vs 2.8< CDAI ≤6 vs CDAI >6 at Week 24.

2. Statistical Hypotheses

The primary objective is to demonstrate that peresolimab is superior to placebo in achieving ACR20 at Week 12. Thus, the null hypothesis to be tested in relation to the primary estimand is:

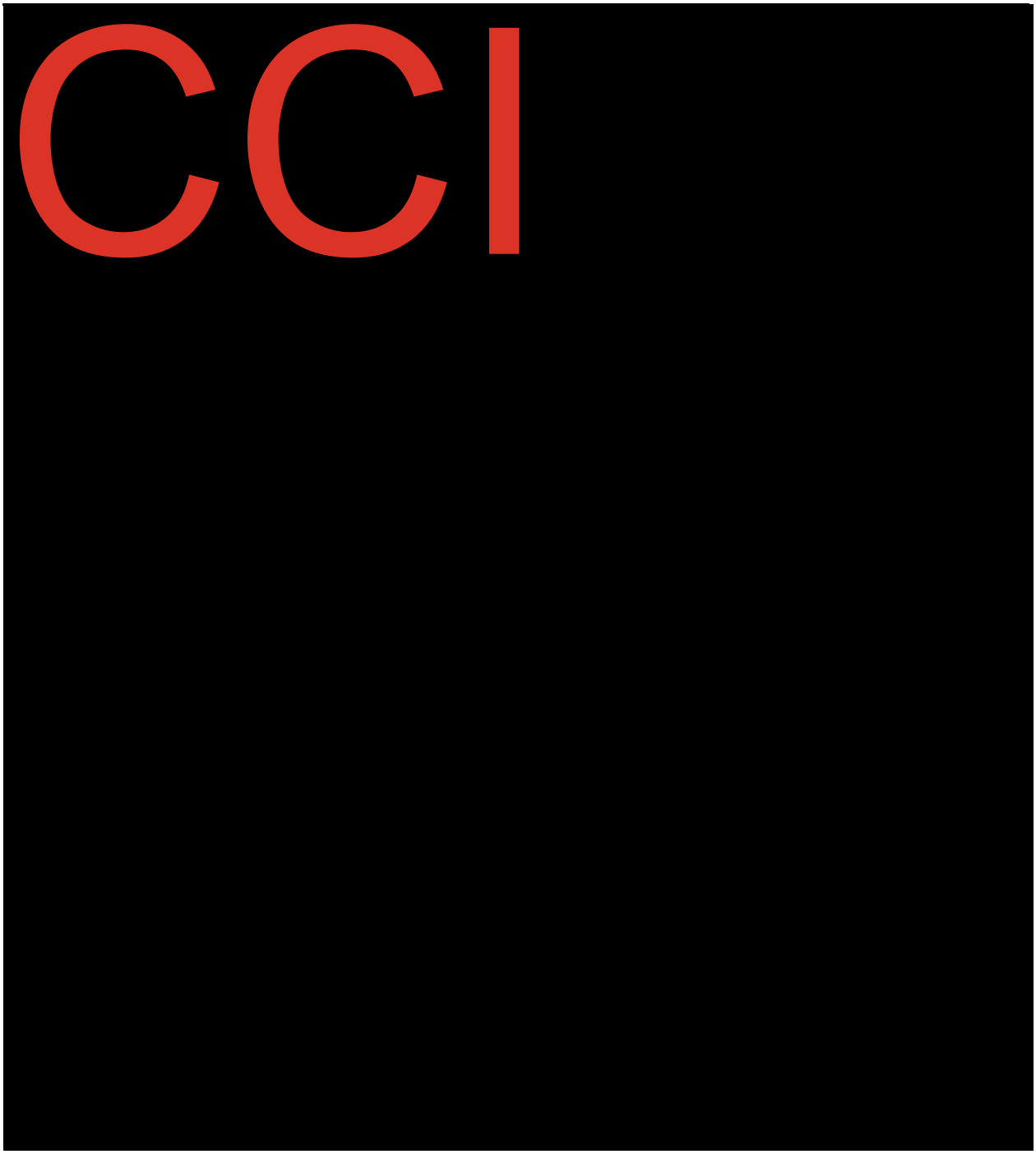
> *There is no difference between peresolimab and placebo when evaluating the proportion of study participants with moderately-to-severely active RA achieving ACR20 at Week 12.*

2.1. Multiplicity Adjustment

Adjustment for multiple comparisons will not be employed in the analysis for this study.

3. Analysis Sets

CCI



Abbreviations: CDAI = Clinical Disease Activity Index; LDA = low disease activity; Q4W = every 4 weeks; Q12W = every 12 weeks; SC = subcutaneous.

Table KDAF.3.1. Treatment Groups and Comparisons for Each Study Period and Analysis Population

Study Period	Analysis Population	Treatment Group	Abbreviation	Inferential Comparison
Double-blind, placebo-controlled treatment period (Period 2)	mITT Population Safety Population	Placebo SC Q4W[a] Peresolimab 100 mg SC Q4W Peresolimab 400 mg SC Q4W Peresolimab 1000 mg SC Q4W	PBO PER100Q4W PER400Q4W PER1000Q4W	PER100Q4W vs. PBO PER400Q4W vs. PBO PER1000Q4W vs. PBO
Double-blind treatment period Weeks 12 to 24 (Period 3A)	Maintenance Population	Placebo SC Q4W to Peresolimab 400 mg SC Q4W Placebo SC Q4W to Peresolimab 1000 mg SC Q4W Peresolimab 100 mg SC Q4W Peresolimab 400 mg SC Q4W Peresolimab 1000 mg SC Q4W	PBO/PER400Q4W PBO/PER1000Q4W PER100Q4W PER400Q4W PER1000Q4W	No Between-Treatment Comparison
Double-blind treatment period Weeks 24 to 60 (Period 3B)	Maintenance Primary Population	Peresolimab 400 mg SC Q4W Peresolimab 400 mg SC Q12W Peresolimab 1000 mg SC Q4W Peresolimab 1000 mg SC Q12W	PER400Q4W PER400Q12W PER1000Q4W PER1000Q12W	No Between-Treatment Comparison
	Maintenance Secondary Population	Peresolimab 400 mg SC Q4W Peresolimab 1000 mg SC Q4W	PER400Q4W PER1000Q4W	No Between-Treatment Comparison
	Maintenance Tertiary Population	Placebo SC Q4W to Peresolimab 400 mg SC Q4W Placebo SC Q4W to Peresolimab 1000 mg SC Q4W Peresolimab 100 mg SC Q4W	PBO/PER400Q4W PBO/PER1000Q4W PER100Q4W	No Between-Treatment Comparison
Combined Periods 2 and 3A (Weeks 0 to 24)	mITT Population	Placebo SC Q4W to Peresolimab 400 mg SC Q4W Placebo SC Q4W to Peresolimab 1000 mg SC Q4W Peresolimab 100 mg SC Q4W Peresolimab 400 mg SC Q4W Peresolimab 1000 mg SC Q4W	PBO/PER400Q4W PBO/PER1000Q4W PER100Q4W PER400Q4W PER1000Q4W	No Between-Treatment Comparison
Treatment Period (combined Periods 2 and 3; Weeks 0 to 60)	Safety Population (for exposure table)	Peresolimab 100 mg SC Q4W Peresolimab 400 mg SC Q4W Peresolimab 1000 mg SC Q4W	PER100Q4W PER400Q4W PER1000Q4W PER400Q4W/Q12W	No Between-Treatment Comparison

Study Period	Analysis Population	Treatment Group	Abbreviation	Inferential Comparison
	Safety Population who Received at Least 1 Dose of Peresolimab	Peresolimab 400 mg SC Q4W to Peresolimab 400 mg SC Q12W Peresolimab 1000 mg SC Q4W to Peresolimab 1000 mg SC Q12W Placebo SC Q4W to Peresolimab 400 mg SC Q4W Placebo SC Q4W to Peresolimab 1000 mg SC Q4W Total (Not applicable for exposure table)	PER1000Q4W/Q12W PBO/PER400Q4W PBO/PER1000Q4W Total	
Posttreatment follow-up period (Period 4)	Follow-up Population	Placebo SC Q4W Peresolimab 100 mg SC Q4W Peresolimab 400 mg SC Q4W Peresolimab 400 mg SC Q12W Peresolimab 1000 mg SC Q4W Peresolimab 1000 mg SC Q12W Total Peresolimab	PBO PER100Q4W PER400Q4W PER400Q12W PER1000Q4W PER1000Q12W Total PER	No Between-Treatment Comparison

Abbreviations: mITT = modified intent-to-treat; PBO = placebo; PER = peresolimab; Q4W every 4 weeks; Q12W = every 12 weeks; SC = subcutaneous.

[a] Participants randomized to either Placebo SC Q4W to Peresolimab 400 mg SC Q4W or Placebo SC Q4W to Peresolimab 1000 mg SC Q4W will be combined as Placebo SC Q4W.

4. Statistical Analyses

4.1. General Considerations

Statistical analysis of this study will be the responsibility of sponsor or its designee. A detailed SAP describing the statistical methodologies will be developed by the sponsor or its designee.

Handling of missing, unused, and spurious data are addressed prospectively in the overall statistical methods described in the protocol and in the SAP, where appropriate. Adjustments to the planned analyses are described in the final clinical study report (CSR).

Baseline for the Treatment Period is defined as the last nonmissing assessment on or prior to the date of the first injection of study treatment for efficacy, unless otherwise specified. In most cases, this will be the measure recorded at Week 0 (Visit 2). For safety analyses, baseline will be defined by period as last nonmissing assessment prior to the first dosing of each period. The Treatment Period starts after the first injection and ends on date of Visit 18 (Week 60) or early discontinuation (ED). Any assessment collected after the first injection is defined as postbaseline for the Treatment Period. If ET visit does not occur, last injection date will be considered as the end of the treatment period. Baseline for safety analysis of the Posttreatment Follow-Up period (Period 4) is defined as the last nonmissing assessment prior to entering Period 4, that is date of Visit 18 (Week 60) or ET.

During this study, data issue was identified for the collection date/time of the electronic clinical outcome assessments (eCOA). Same analysis visit was assigned to eCOA assessment with different collection date than the corresponding visit date. Because of this, study team came up with a method to handle this data issue, which is the analysis visit numbers for electronic clinical outcome assessments (eCOA) data will be assigned based on the actual visit date recorded in the electronic case report form (eCRF). If the eCOA assessment date is within +/- 3 days window of the actual visit, then the analysis visit number of the eCOA assessment will be assigned as the actual visit number. If the eCOA assessment date is outside +/- 3 days window of all the actual visits, then the analysis visit of the eCOA assessment will be considered as unscheduled visit. The eCOA data will be used for the analysis only when the analysis visit number of the eCOA assessment is identical to the visit number captured in the eCOA system.

For analyses in Period 2, participants randomized to either Placebo subcutaneous (SC) Q4W to Peresolimab 400 mg SC Q4W or Placebo SC Q4W to Peresolimab 1000 mg SC Q4W will be combined as Placebo SC Q4W.

Dichotomous efficacy endpoints including the primary endpoint during Period 2 will be analyzed using a CCI [redacted]. For each treatment comparison vs. Placebo SC Q4W during Period 2, an estimate of the treatment difference with corresponding Wald 95% confidence interval (CI), the odds ratio with corresponding Wald 95% CI and p-value will be presented. Each treatment group will be compared to Placebo SC Q4W.

Continuous efficacy endpoints, which are planned to be collected at >1 postbaseline timepoints during Period 2 according to the SoA, will be analyzed using mixed-effect model for repeated measures (MMRM) analysis. The MMRM model will include:

- treatment, stratification factors, baseline value, visit and treatment-by-visit interaction in the model as fixed factors, and
- participant as a random factor.

The least squares (LS) mean for each treatment group along with the estimate of the difference between each treatment group with Placebo SC Q4W, standard error (SE), p-value, and the 95% CIs will be reported at each visit along with p-values.

The covariance structure to model the within-participant errors will be unstructured. If the unstructured covariance matrix results in a lack of convergence, the heterogeneous Toeplitz covariance structure, followed by the CCI [redacted], will be used.

The CCI [redacted] will be used to estimate the denominator degrees of freedom. Type III sums of squares for the LS means will be used for the statistical comparison. The 95% CI will also be reported.

When continuous efficacy endpoints are analyzed using CCI [redacted], the model will include treatment group, stratification factors, and baseline value as covariates. Type III sums of squares for the LS means will be used for statistical comparison between each treatment group and Placebo SC Q4W. The LS mean difference, SE, p-value, and 95% CI, unless otherwise specified, will also be reported.

Dose response modeling will be performed on the primary, and other core efficacy endpoints, if deemed necessary, to assist in dose selection decisions. More details will be provided in Section 4.2.

Data collected after initiation of rescue or prohibited medication use will be handled according to the specified estimand strategy for efficacy analysis. Such data, unless otherwise noted, will be excluded from efficacy analyses, considered as missing, and handled according to a missing data imputation method as specified.

Missing data will be handled separately from estimand strategy. After intercurrent events related to study intervention are accounted for according to an estimand strategy specified in Section 1.1, any remaining missing data will be handled using the following missing data imputation methods:

- Nonresponder imputation (NRI):

  When NRI is applied, participants with missing data will be considered as nonresponders regardless of whether they have at least 1 postbaseline data.

- LOCF:

When LOCF is applied, the last nonmissing observation up to Week 24 will be carried to the corresponding time point for evaluation. For participants who initiated rescue therapy, the last nonmissing observation prior to the initiation of rescue therapy will be carried forward. For randomized participants without any postbaseline observation, missing data will not be imputed.

4.2. Bayesian Model Averaging

A Bayesian model averaging (BMA) approach will be used to estimate the dose response relationship. This BMA approach is the Bayesian analog of the multiple comparisons procedure - modeling (MCP-MOD) methodology (Bretz et al. 2005), and "the Qualification of the MCP-Mod procedure" (FDA 2015) is supportive in the use of MCP-MOD or BMA to assist in dose selection decisions.

BMA is a general mixture distribution, where each mixture component is a different parametric model. Prior weights are placed on each model and the posterior model weights are updated based on how well each model fits the data. Let $\mu(d)$ represent the mean of the dose response curve at dose d, $y = \{y_1, \ldots, y_n\}$ be the observed data, and $m \in \{1, \ldots, M\}$ be an index on the M parametric models described below. Then the posterior of the dose response curve, $\mu(d)$, of the BMA model is

$$p(\mu(d) \mid y) = \sum_{m=1}^{M} p(\mu(d) \mid y, m)\, p(m \mid y)$$

$$p(m \mid y) = \frac{p(y \mid m)p(m)}{\sum_{m^*} p(y \mid m^*)p(m^*)}$$

where $p(\mu(d) \mid y, m)$ is the posterior mean dose response curve from model m, $p(m \mid y)$ is the posterior weight of model m, $p(y \mid m)$ is the marginal likelihood of the data under model m, and $p(m)$ is the prior weight assigned to model m. In cases where $p(y \mid m)$ is difficult to compute, Gould (2019) proposes using the observed data's fit to the posterior predictive distribution as a surrogate in calculating the posterior weights; this is the approach used in this analysis.

4.2.1.1. BMA Analysis and Reporting

CCI [redacted]

BMA analysis will be summarized by dose and provide:

- Observed response
- BMA estimated
  - Mean response
  - SE and standard deviation of the mean response
  - 2.5% and 97.5% quantiles
  - Summaries of posterior probability of LY-PBO treatment effect, that is, Prob(LY-PBO > EOI) where EOI is effect of interest

- Plot that includes observed response, BMA estimated response, and 2.5% and 97.5% posterior quantiles

BMA analysis will be summarized by component and model parameter and provide:

- Posterior mean
- SE and standard deviation of the posterior mean
- 2.5% and 97.5% posterior quantiles of the posterior distribution
- Convergence diagnostics
- Effective sample size
- Prior distributions
- Plot of the fitted BMA
- Plot of the BMA components

See Appendix 9 (Section 6.9) for execution details.

4.3. Participant Dispositions

A detailed description of participant disposition will be provided, including a summary of the number and percentage of participants entered into the study and randomized, the number and percentage of participants who complete the study or discontinue, both overall and by reason for discontinuation, and the frequency and percentage of patients who discontinued study treatment. A summary of important protocol deviations will be provided.

4.4. Efficacy Analyses

Table KDAF.4.1 includes the description and derivation of efficacy outcomes, including the primary and secondary efficacy outcomes.

Table KDAF.4.2 provides the detailed analysis plans, including estimand, analysis type, analysis method, missing data imputation, analysis population, timepoint, and treatment comparisons for the efficacy analyses.

4.4.1. Primary Endpoint and Analysis Methodology

The primary endpoint is the ACR20 response rate at Week 12.

ACR20 is defined as at least 20% improvement from baseline in both TJC (68 joint count; 0 to 68) and SJC (66 joint count; 0 to 66), and at least 20% improvement from baseline in at least 3 of the following assessments:

- Patient's Assessment of Arthritis Pain (VAS; 0 to 100 mm)
- Patient's Global Assessment of Disease Activity (VAS; 0 to 100 mm)
- Physician's Global Assessment of Disease Activity (VAS; 0 to 100 mm)
- Patient's Assessment of Physical Function as measured by the HAQ-DI (0 to 3 by 0.125)
- Acute Phase reactant as measured by high-sensitivity C-reactive protein (hsCRP) (mg/dL)

Refer to the algorithm in Appendix 6 (Section 6.6) to calculated ACR response.

The primary objective of this study is to test the null hypothesis that there is no difference in achieving ACR20 at Week 12 between peresolimab and placebo in patients with moderately-to-severely active RA.

CCI [redacted]

Any missing data (for example, missing due to missing a visit) will be imputed using NRI for the primary analysis.

4.4.2. Secondary Endpoints and Analysis Methodology

The secondary endpoints at Week 12 in Period 2 include:

- Proportion of participants achieving:
  - ACR50
  - ACR70
  - Remission or LDA in terms of DAS28-hsCRP
  - Remission or LDA in terms of DAS28-ESR
  - Remission or LDA in terms of SDAI, and
  - Remission or LDA in terms of CDAI
- Mean change from baseline in:
  - DAS28-hsCRP
  - SDAI
  - CDAI
  - TJC68
  - SJC66, and
  - Physician's Global Assessment of Disease Activity (VAS)
  - Patient's Global Assessment of Disease Activity (VAS)
  - Patient's Assessment of Arthritis Pain (VAS), and
  - patient's assessment of physical function using HAQ-DI
  - Duration of morning joint stiffness
  - Severity of morning joint stiffness
  - FACIT-F scores
  - Short Form-36 health survey acute form (SF-36) domains
  - SF-36 Physical Component Summary, and
  - SF-36 Mental Component Summary.

The secondary endpoints at Week 24 and Week 60 in Period 3 include:

- Proportion of participants achieving:
  - ACR20
  - ACR50
  - ACR70
  - Remission or LDA in terms of DAS28-hsCRP
  - Remission or LDA in terms of DAS28-ESR
  - Remission or LDA in terms of SDAI, and
  - Remission or LDA in terms of CDAI
- Mean change from baseline in:
  - DAS28-hsCRP
  - SDAI
  - CDAI
  - TJC68
  - SJC66, and
  - Physician's Global Assessment of Disease Activity (VAS).

4.4.2.1. Main Analytical Approach

4.4.2.1.1. ACR50 and ACR70

ACR50 and ACR70 are calculated in the same manner as ACR20 but are based on 50% and 70% improvement, respectively.

Summaries and analyses similar to those described for ACR20 in Section 4.4.1 will be conducted for ACR50 and ACR70. Any missing data (for example, missing due to missing a visit) will be imputed using NRI.

4.4.2.1.2. DAS28-hsCRP

The change from baseline of DAS28-hsCRP at Week 12 for the peresolimab treatment group will be compared with placebo.

An MMRM model, as described in Section 4.1, will be used to test the treatment difference of each peresolimab treatment group versus placebo. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will not be imputed.

4.4.2.1.3. SDAI

The change from baseline of SDAI at Week 12 for the peresolimab treatment group will be compared with placebo.

A MMRM model, as described in Section 4.1, will be used to test the treatment difference of each peresolimab treatment group versus placebo. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will not be imputed.

4.4.2.1.4. CDAI

The change from baseline of CDAI at Week 12 for the peresolimab treatment group will be compared with placebo.

A MMRM model, as described in Section 4.1, will be used to test the treatment difference of each peresolimab treatment group versus placebo. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will not be imputed.

4.4.2.1.5. ACR Core Set Values

The change from baseline of the following individual components at Week 12 of the ACR core set will be compared with placebo:

- TJC (68 joint count; 0 to 68)
- SJC (66 joint count; 0 to 68)
- Patient's Assessment of Pain (VAS; 0 to 100 mm)
- Patient's Global Assessment of Disease Activity (VAS; 0 to 100 mm)
- Physician's Global Assessment of Disease Activity (VAS; 0 to 100 mm)
- Patient's Assessment of Physical Function as measured by the HAQ-DI (0 to 3 by 0.125)
- Acute Phase reactant as measured by hsCRP (mg/dL)

A MMRM model, as described in Section 4.1, will be used to test the treatment difference of each peresolimab treatment group versus placebo. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will not be imputed.

4.4.2.1.6. DAS28-hsCRP Remission and Low Disease Activity

DAS28-hsCRP remission and LDA at Week 12 for the peresolimab treatment group will be compared with placebo.

A logistic regression model will be used, as described in Section 4.1, to test the treatment difference of each peresolimab treatment group versus placebo. A composite strategy will be used to handle for intercurrent events. Any missing data (for example, missing due to missing a visit) will be imputed using NRI.

4.4.2.1.7. DAS28-ESR Remission and Low Disease Activity

DAS28-ESR remission and LDA at Week 12 for the peresolimab treatment group will be compared with placebo.

A logistic regression model will be used, as described in Section 4.1, to test the treatment difference of each peresolimab treatment group versus placebo. A composite strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will be imputed using NRI.

4.4.2.1.8. SDAI Remission and Low Disease Activity

SDAI remission and LDA at Week 12 for the peresolimab treatment group will be compared with placebo.

A logistic regression model will be used, as described in Section 4.1, to test the treatment difference of each peresolimab treatment group versus placebo. A composite strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will be imputed using NRI.

4.4.2.1.9. CDAI Remission and Low Disease Activity

CDAI remission and LDA at Week 12 for the peresolimab treatment group will be compared with placebo.

A logistic regression model will be used, as described in Section 4.1, to test the treatment difference of each peresolimab treatment group versus placebo. A composite strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will be imputed using NRI.

4.4.2.1.10. Duration and Severity of Morning Joint Stiffness

The change from baseline of the duration and severity of morning joint stiffness at Week 12 for the peresolimab treatment group will be compared with placebo.

A MMRM model, as described in Section 4.1, will be used to test the treatment difference of each peresolimab treatment group versus placebo. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will not be imputed.

4.4.2.1.11. FACIT-F

The change from baseline of FACIT-F at Week 12 for the peresolimab treatment group will be compared with placebo.

A MMRM model, as described in Section 4.1, will be used to test the treatment difference of each peresolimab treatment group versus placebo. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will not be imputed.

4.4.2.1.12. SF-36 Domain and Component Summary

The change from baseline of Short Form-36 version 2 health survey acute form (SF-36v2) acute for the peresolimab treatment group will be compared with placebo.

An ANCOVA model, as described in Section 4.1, will be used to test the treatment difference of each peresolimab treatment group versus placebo. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will be imputed using LOCF.

4.4.3. Supportive Secondary Endpoints

4.4.3.1. Definition of Endpoints

The supportive secondary endpoints include:

- Proportion of participants achieving LDA or remission at Weeks 24 and 60 for
  - DAS28-hsCRP

  - DAS28-ESR
  - SDAI, and
  - CDAI
- Proportion of participants achieving ACR20, ACR50, or ACR70 at Weeks 24 and 60
- Change from baseline at Weeks 24 and 60 for ACR core set values
  - TJC68
  - SJC66, and
  - Physician's Global Assessment of Disease Activity (VAS)
- Change from baseline at Weeks 24 and 60 for mean
  - DAS28-hsCRP
  - SDAI, and
  - CDAI
- Observed trough drug concentration at Week 12

4.4.3.2. Main Analytic Approach

For continuous endpoints, summary tables at assessed time points and the change from the baseline assessment will be presented by treatment groups. Observed mean, standard deviation, minimum, median, and maximum will be reported. A hypothetical strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will be imputed with LOCF.

For dichotomous endpoints, summary tables at assessed time points will be presented by treatment groups. The number of observed responses, the rate of observed response, and 95% CI will be reported. A composite strategy will be used to handle intercurrent events. Any missing data (for example, missing due to missing a visit) will be inputted using NRI.

4.5. Exploratory Endpoints Analysis

CCI


CCI


CCI

Table KDAF.4.1. Description and Derivation of Efficacy and Patient Reported Outcome Measures

Measure	Description	Variable	Derivation/Comment	Definition of Missing
American College of Rheumatology (ACR)	ACR score is a composite of clinical, laboratory, and functional measures in RA to assess relief of signs and symptoms.	ACR20/50/70	See Appendix 6 (Section 6.6) for details.	See Appendix 6 (Section 6.6) for details.
Disease Activity Score – high-sensitivity C-reactive protein (DAS28-hsCRP)	The DAS28-CRP is a measure of disease activity in 28 joints that consists of a composite numerical score utilizing the following variables: TJC, SJC, high-sensitivity C-reactive protein (hsCRP) (measured in mg/L), and Patient's Global Assessment of Disease Activity (PatGA) recorded by patients on a 0- to 100-mm VAS. For DAS28-hsCRP, the 28 joints to be examined and assessed as tender or not tender for TJC (TJC 28) and as swollen or not swollen for SJC (SJC 28) are a subset of those assessed for the TJC and SJC and include 14 joints on each side of the patient's body: the 2 shoulders, the 2 elbows, the 2 wrists, the 10 metacarpophalangeal joints, the 2 interphalangeal joints of the thumb, the 8 proximal interphalangeal joints, and the 2 knees. (Vander Cruyssen et al. 2005)	DAS28-hsCRP	$\text{DAS28-hsCRP} = 0.56\sqrt{\text{TJC28}} + 0.28\sqrt{\text{SJC28}} + 0.36\ln(\text{hsCRP} + 1) + 0.014\text{PatGA} + 0.96$	If 1 or more variables are missing, then set to missing.
		Change from baseline in DAS28-hsCRP	Calculated as: Observed DAS28-hsCRP score – baseline DAS28-hsCRP	Missing if baseline or observed value are missing.
		DAS28-hsCRP Remission	Remission is defined as observed DAS28-hsCRP score of <2.6.	Missing if observed is missing.
		DAS28-hsCRP Low Disease Activity (LDA)	DAS28-hsCRP LDA is defined as observed DAS28-hsCRP score of ≤3.2.	Missing if observed is missing.
Disease Activity Score – erythrocyte sedimentation rate (DAS28-ESR)	Similar to DAS28-hsCRP, except erythrocyte sedimentation rate (ESR) is used in replacement of hsCRP.	DAS28-ESR	$\text{DAS28-ESR} = 0.56\sqrt{\text{TJC28}} + 0.28\sqrt{\text{SJC28}} + 0.36\ln(\text{ESR} + 1) + 0.014\text{PatGA} + 0.96$	If 1 or more variables are missing, then set to missing.
		Change from baseline in DAS28-ESR	Calculated as: Observed DAS28-ESR score – baseline DAS28-ESR	Missing if baseline or observed value are missing.
		DAS28-ESR Remission	Remission is defined as observed DAS28-ESR score of <2.6.	Missing if observed is missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
		DAS28-ESR LDA	DAS28-hsCRP LDA is defined as observed DAS28-ESR score of ≤3.2.	Missing if observed is missing.
Simplified Disease Activity Index (SDAI)	SDAI is a tool for measurement of disease activity in RA that integrates measures of physical examination, acute phase response, patient self-assessment, and evaluator assessment. (Aletaha and Smolen 2005)	SDAI score	$SDAI = TJC28 + SJC28 + hsCRP/10 + PGA + PatGA,$ Note: Acute Phase reactant as measured by hsCRP (mg/dL) ranging from 0.1 to 10	If 1 or more variables are missing, then set to missing.
		Change from baseline in SDAI	Calculated as: Observed SDAI score – baseline SDAI	Missing if baseline or observed value are missing.
		SDAI Remission	Remission is defined as observed SDAI score of <3.3	Missing if observed is missing.
		SDAI LDA	LDA is defined as observed SDAI score ≤11	Missing if observed is missing.
Clinical Disease Activity Index (CDAI)	CDAI is similar to the SDAI, but it allows for immediate scoring because it does not use a laboratory result. (Aletaha and Smolen 2005)	CDAI score	$CDAI = TJC28 + SJC2 + PGA + PatGA$	If 1 or more variables are missing, then set to missing.
		Change from baseline in CDAI	Calculated as: Observed CDAI score – baseline CDAI	Missing if baseline or observed value are missing.
		CDAI Remission	Remission is defined as observed CDAI score of ≤2.8	Missing if observed is missing.
		CDAI LDA	LDA is defined as observed CDAI score ≤10	Missing if observed is missing.
Tender Joint Count-68 joint counts (TJC68)	The number of tender and painful joints will be determined by examination of 68 joints (34 joints on each side of the patient's body). The 68 joints to be assessed and classified as tender or not tender.	TJC68 score	TJC68 is calculated as: rounding up (number of tender joints)/(number of evaluable joints)×68; See Appendix 7 (Section 6.7) for details.	If more than half of the individual joint scores are nonevaluable, total score missing.
		Change from baseline in TJC68	Calculated as: Observed TJC68 – baseline TJC68	Missing if baseline or observed value are missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
		Percent improvement from baseline in TJC68	Calculated as: 100 × (baseline TJC68 – observed TJC68)/baseline TJC68	Missing if baseline or observed value are missing.
Tender Joint Count – 28 joint counts (TJC28)	The number of tender and painful joints will be determined by examination of 28 joints (14 joints on each side of the patient's body). The 28 joints to be assessed and classified as tender or not tender.	TJC28 score	TJC28 is calculated as: rounding up (number of tender joints)/(number of evaluable joints)×28 See Appendix 7 (Section 6.7) for details.	If more than half of the individual joint scores are nonevaluable, total score missing.
		Change from baseline in TJC28	Calculated as: Observed TJC28 – baseline TJC28	Missing if baseline or observed value are missing.
		Percent improvement from baseline in TJC28	Calculated as: 100 × (baseline TJC28 – observed TJC28)/baseline TJC28	Missing if baseline or observed value are missing.
Swollen Joint Count – 66 joint counts (SJC66)	The number of swollen joints will be determined by examination of 66 joints (33 joints on each side of the patient's body). The 66 joints to be assessed and classified as swollen or not swollen.	SJC66 score	SJC66 is calculated as: rounding up (number of swollen joints)/(number of evaluable joints)×66 See Appendix 7 (Section 6.7) for details.	If more than half of the individual joint scores are nonevaluable, total score missing.
		Change from baseline in SJC66	Calculated as: Observed SJC66 – baseline SJC66	Missing if baseline or observed value are missing.
		Percent improvement from baseline in SJC66	Calculated as: 100 × (baseline SJC66 – observed SJC66)/baseline SJC66	Missing if baseline or observed value are missing.
Swollen Joint Count – 28 joint counts (SJC28)	The number of swollen joints will be determined by examination of 28 joints (14 joints on each side of the patient's body). The 28 joints to be assessed and classified as swollen or not swollen.	SJC28 score	SJC 28 is calculated as: rounding up (number of swollen joints)/(number of evaluable joints)×28 See Appendix 7 (Section 6.7) for details.	If more than half of the individual joint scores are nonevaluable, total score missing.
		Change from baseline in SJC28	Calculated as: Observed SJC28 – baseline SJC28	Missing if baseline or observed value are missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
		Percent improvement from baseline in SJC28	Calculated as: 100 × (baseline SJC28 – observed SJC28)/baseline SJC28	Missing if baseline or observed value are missing.
Physician's Global Assessment of Disease Activity (PGA) Visual Analog Scale (VAS)	Overall assessment of the severity of the patient's current RA activity using a 100-mm horizontal VAS. The investigator making the assessment must be a rheumatologist or medically qualified physician	PGA VAS score	No derivation; ranges from 0 to 100 mm, where 0 represents no disease activity and 100 represents extremely active disease activity.	Missing if observed value is missing.
		Change from baseline in PGA VAS	Calculated as: Observed PGA – baseline PGA	Missing if baseline or observed value are missing.
		Percent improvement from baseline in PGA VAS	Calculated as: 100 × (baseline PGA – observed PGA)/baseline PGA	Missing if baseline or observed value are missing.
PatGA VAS	Assesses the patient's overall assessment of his or her RA activity by marking a 100-mm horizontal VAS. (Nikiphorou et al. 2016)	PatGA VAS score	No derivation; ranges from 0 to 100 mm, where 0 represents no disease activity and 100 represents extremely active disease activity.	Missing if observed value is missing.
		Change from baseline in PatGA VAS	Calculated as: Observed PatGA – baseline PatGA	Missing if baseline or observed value are missing.
		Percent improvement from baseline in PatGA VAS	Calculated as: 100 × (baseline PatGA – observed PatGA)/baseline PatGA	Missing if baseline or observed value are missing.
Patient's Assessment of Arthritis Pain (Pain) VAS	Assesses the patient's current level of joint pain by marking a vertical tick on a 100-mm horizontal VAS.	Pain VAS score	No derivation; ranges from 0 to 100 mm, where 0 represents no joint pain and 100 represents the worse joint pain.	Missing if observed value is missing.
		Change from baseline in Pain VAS	Calculated as: Observed Pain – baseline Pain	Missing if baseline or observed value are missing.
		Percent improvement from baseline in Pain VAS	Calculated as: 100 × (baseline Pain – observed Pain)/baseline Pain	Missing if baseline or observed value are missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
Patient's Assessment of Physical Function Health Assessment Questionnaire – Disability Index (HAQ-DI)	Patient-reported standardized questionnaire that is commonly used in RA to measure disease-associated disability (assessment of physical function). It consists of 24 questions referring to 8 domains: dressing/grooming, arising, eating, walking, hygiene, reach, grip, and other daily activities. (Fries et al. 1980, 1982; Fries 1983; Ramey et al. 1996)	HAQ-DI score	Sum of the highest subcategory scores within 8 categories and adjust for aids/devices and/or help from another person. It ranges from 0 to 3 by 0.125. See Appendix 7 (Section 6.7) for details.	Missing if <6 categories completed.
		Change from baseline in HAQ-DI	Calculated as: Observed HAQ-DI – baseline HAQ-DI	Missing if baseline or observed value are missing.
		Percent improvement from baseline in HAQ-DI	Calculated as: 100 × (baseline HAQ-DI – observed HAQ-DI)/baseline HAQ-DI	Missing if baseline or observed value are missing.
		HAQ-DI Minimal Clinically Important Difference (MCID)	MCID is defined as baseline HAQ-DI – observed HAQ-DI ≥0.22 Note: participants with HAQ-DI score <0.22 at baseline will be excluded from the analysis set. Additional threshold may be considered if necessary.	Missing if baseline or observed value are missing.
hsCRP (mg/dL)	The ACR core set measure and DAS28-hsCRP core set measure of acute phase reactant. Assayed by Lilly-designated laboratory.	hsCRP	No derivation as raw laboratory result	Missing if missing.
		Change from baseline in hsCRP	Calculated as: Observed hsCRP – baseline hsCRP	Missing if baseline or observed value are missing.
		Percent improvement from baseline in hsCRP	Calculated as: 100 × (baseline hsCRP – observed hsCRP)/baseline hsCRP	Missing if baseline or observed value are missing.
ESR (mm/h)	Laboratory measure in DAS28-ESR. Assayed and evaluated locally.	ESR	No derivation as raw laboratory result	Missing if missing.
		Change from baseline in ESR	Calculated as: Observed ESR – baseline ESR	Missing if baseline or observed value are missing.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
		Percent improvement from baseline in ESR	Calculated as: 100 × (baseline ESR – observed ESR)/baseline ESR	Missing if baseline or observed value are missing.
Duration and severity of Morning Joint Stiffness	The duration of morning joint stiffness is a patient-administered item that allows for the patients to enter the length of time in minutes that their morning joint stiffness lasted on the day prior to that visit using an electronic patient-reported outcomes (ePRO) tablet. Durations recorded as longer than 12 hours (720 minutes) will be truncated to 720 minutes for analysis. The severity of morning joint stiffness is a single item, 11-point horizontal scale that captures the severity of morning joint stiffness using a scale from 0 to 10, where 0 implies no joint stiffness, and 10 implies joint stiffness as bad as you can imagine. The worst level of joint stiffness in the last 7 days will be used for analysis.	Duration of morning joint stiffness	No derivation	Missing if missing.
		Change from baseline in duration of morning joint stiffness	Calculated as: Observed duration – baseline duration	Missing if baseline or observed value are missing.
		Severity of morning joint stiffness	No derivation	Missing if missing.
		Change from baseline in severity of morning joint stiffness	Calculated as: Observed duration – baseline duration	Missing if baseline or observed value are missing.
Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-F)	The FACIT-F scale (Cella and Webster 1997) is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning. The FACIT-F uses 0 (“not at all”) to 4 (“very much”) numeric rating scales to assess fatigue and its impact in the past 7 days. Scores range from 0 to 52 with higher scores indicating less fatigue. (Cella 1997)	FACIT-F score	Sum of 13 symptom-specific items. It ranges from 0 to 52 with 0 being the worst possible score. See Appendix 7 (Section 6.7) for details.	Missing if >6 items are missing.
		Change from baseline in FACIT-F	Calculated as: Observed FACIT-F – baseline FACIT-F	Missing if baseline or observed value are missing.
		8 associated domain scores:	Refer to the manual “How to Score Version 2 of the SF-36 Health Survey” (Ware et al. 2000)	To be added.

Measure	Description	Variable	Derivation/Comment	Definition of Missing
Short Form-36 version 2 health survey acute form (SF-36v2)	The SF-36v2 Acute measure is a subjective, generic, health-related quality of life instrument that is patient-reported and consists of 36 questions covering 8 health domains: physical functioning, bodily pain, role limitations due to physical problems, role limitations due to emotional problems, general health perceptions, mental health, social function, and vitality. Each domain is scored by summing the individual items and transforming the scores into a 0 to 100 scale with higher scores indicating better health status or functioning. In addition, 2 summary scores, the physical component score (PCS) and the mental component score (MCS), will be evaluated based on the 8 SF-36v2 Acute domains (McHorney et al. 1993; Ware et al. 2007; Ware and Sherbourne 1992; Maruish 2011).	• Physical Functioning • Role Physical • Bodily Pain • General Health • Vitality • Social Functioning • Role Emotional • Mental Health 2 component Scores: • MCS Score • PCS Score		
		Change from baseline in 8 associated domain scores and 2 component scores	Calculated as: Observed value – baseline value	Missing if baseline or observed value are missing.

Abbreviations: PGA = physician's global assessment; PatGA = patient's global assessment; RA = rheumatoid arthritis; SF-36 = Short Form-36 health survey acute form; SJC = swollen joint count; TJC = tender joint count.

Table KDAF.4.2. Description of Efficacy and Patient Reported Outcome Analyses

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
American College of Rheumatology (ACR)	ACR20	Composite	Logistic with NRI	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Primary analysis is at Week 12. CCI [redacted]
		Composite	Descriptive statistics with NRI	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI [redacted]
		Composite	Descriptive statistics with NRI	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI [redacted]
	ACR50/70	Composite	Logistic with NRI	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI [redacted]
		Composite	Descriptive statistics with NRI	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI [redacted]
		Composite	Descriptive statistics with NRI	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI [redacted]

Approved on 01 Dec 2024 GMT

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
DAS28-hsCRP	Change from baseline (and/or percent improvement from baseline)	Hypothetical	MMRM with no missing data imputation	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI [redacted]
		Hypothetical	Descriptive statistics with LOCF	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI [redacted]
		Hypothetical	Descriptive statistics with LOCF	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI [redacted]
	LDA; Remission	Composite	Logistic with NRI	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI [redacted]
		Composite	Descriptive statistics with NRI	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI [redacted]
		Composite	Descriptive statistics with NRI	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI [redacted]

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
CCI						
DAS28-ESR	LDA; Remission	Composite	Logistic with NRI	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI
		Composite	Descriptive statistics with NRI	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI
		Composite	Descriptive statistics with NRI	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
CCI						
SDAI; CDAI	Change from baseline (and/or percent improvement from baseline)	Hypothetical	MMRM with no missing data imputation	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI
		Hypothetical	Descriptive statistics with LOCF	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI
		Hypothetical	Descriptive statistics with LOCF	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
	LDA; Remission	Composite	Logistic with NRI	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI
		Composite	Descriptive statistics with NRI	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI
		Composite	Descriptive statistics with NRI	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI

CCI

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
TJC68; SJC66; PGA VAS; PatGA VAS; Pain VAS	Change from baseline (and/or percent improvement from baseline)	Hypothetical	MMRM with no missing data imputation	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI [redacted]
		Hypothetical	Descriptive statistics with LOCF	Maintenance	No between-treatment comparison; At Week 24 and all other scheduled visits in Period 3A	Secondary analysis is at Week 24. CCI [redacted]
		Hypothetical	Descriptive statistics with LOCF	Maintenance Primary; Maintenance Secondary; Maintenance Tertiary	No between-treatment comparison; At Week 60 and all other scheduled visits in Period 3B	Secondary analysis is at Week 60. CCI [redacted]
HAQ-DI	Change from baseline (and/or percent improvement from baseline)	Hypothetical	MMRM with no missing data imputation	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI [redacted]
CCI [redacted]						

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
CCI						
FACIT-F	Change from baseline	Hypothetical	MMRM with no missing data imputation	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI
CCI						

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
CCI						
Duration and Severity of Morning Joint Stiffness	Change from baseline	Hypothetical	MMRM with no missing data imputation	mITT	PER100Q4W vs. PBO; PER400Q4W vs. PBO; PER1000Q4W vs. PBO; At Week 12 and all other scheduled visits in Period 2	Secondary analysis is at Week 12. CCI
CCI						

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
CCI						

Measure	Variable	Estimand	Analysis Method	Population	Comparison and Timepoint	Analysis Type
CCI						

Abbreviations: BOCF = baseline observation carried forward; CDAI = Clinical Disease Activity Index; DAS28-ESR = Disease Activity Score – erythrocyte sedimentation rate; DAS28-hsCRP = Disease Activity Score – high-sensitivity C-reactive protein; ESR = erythrocyte sedimentation rate; FACIT-F = Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI = Health Assessment Questionnaire – Disability Index; hsCRP = high-sensitivity C-reactive protein; LDA = low disease activity; LOCF = last observation carried forward; MCID = Minimal Clinically Important Difference; mITT = modified intent-to-treat; MMRM = mixed-effect model for repeated measures; NRI = nonresponder imputation; Pain = Patient's Assessment of Arthritis Pain; PatGA = Patient's Global Assessment of Disease Activity; PBO = placebo; PER = Peresolimab; PGA = Physician's Global Assessment; Q4W = every 4 weeks; Q12W = every 12 weeks; RAMRIS = Rheumatoid Arthritis Magnetic Resonance Imaging Score; SDAI = Simplified Disease Activity Index; SJC66 = Swollen Joint Count – 66 joint counts; TJC68 = Tender Joint Count- 68 joint counts; VAS = visual analog scale.

4.6. Safety Analyses

All safety data will be descriptively summarized by treatment groups and by treatment periods. Safety data for Period 2, Period 3A, and Period 3B will be analyzed based on the Safety Population, Maintenance Population, and Maintenance Primary/Secondary/Tertiary Population, respectively, as defined in Section 3. For Period 3B based on Maintenance Primary/Secondary/Tertiary Population, the treatment groups will be combined and not be separated by populations. Patients will be summarized according to the treatment to which they were assigned at Week 24. Safety summaries will be presented by treatment arm as defined in Table 3.1. The safety analyses include adverse events (AEs), safety in special groups and circumstances, including, but not limited to, adverse events of special interest, laboratory analytes, and vital signs.

4.6.1. Extent of Exposure

Duration of exposure to study drug during the double-blind treatment period will be summarized by treatment arm and provided as listing. Duration of exposure on treatment will be calculated as the date of last dose of study drug (or date of discontinuation) minus the date of first dose of study drug plus 1 day. Total patient-years (PYs) of exposure will be reported for each treatment group for overall duration of exposure. Descriptive statistics (n, mean, SD, minimum, maximum, first quartile, median, third quartile, and maximum) will be provided for patient-days of exposure and the frequency of patient falling into different exposure ranges will be summarized. Exposure ranges are as follows:

- >0, ≥12 weeks, ≥24 weeks, and ≥60 weeks
- >0 to <12 weeks, ≥12 weeks to <24 weeks, ≥24 weeks to <60 weeks, and ≥60 weeks

Overall exposure will be summarized in total PYs, which will be calculated as follows:

Exposure in patient years = Sum of duration of exposure (for all patients in treatment arm) / 365.25.

No inferential analysis for comparison between treatment arms will be performed.

4.6.2. Adverse Events

AEs are recorded in the eCRF. Each AE will be coded to System Organ Class (SOC) and Preferred Term (PT), using the *Medical Dictionary for Regulatory Activities* (MedDRA) version that is current at the time of database lock. Severity of AEs is recorded as mild, moderate, or severe.

A treatment-emergent adverse event (TEAE) is defined as an event that first occurred or worsened in severity after baseline and on or prior to the earliest of the visit study drug disposition dates or the last visit date during the treatment period, whichever occurred first, and up to 50 days after study treatment discontinuation. Baseline for Period 2 is defined as all preexisting conditions recorded at Visit 1 and any AEs recorded before the first dose of study intervention. Baselines for other periods are defined as unresolved AEs just prior to the first dose of each period. The MedDRA Lowest Level Term (LLT) will be used in defining which events

are treatment-emergent. The maximum severity for each LLT will be used as baseline. If an event is preexisting during the baseline period for Period 2, but it has missing severity, and the event persists during the treatment period or up to 50 days after treatment discontinuation, then the baseline severity will be considered mild for determining any postbaseline treatment-emergence (that is, the event is treatment-emergent unless the severity is coded mild at postbaseline). If an event occurring postbaseline has a missing severity rating, then the event is considered treatment-emergent. Should there be insufficient data for an AE start date to make this comparison (for example, the AE start year is the same as the treatment start year, but the AE start month and day are missing), the AE will be considered treatment-emergent. For events occurring on the day of the first dose of study treatment, Week 0 (V2) in most cases, the events will be assumed to occur posttreatment and considered as TEAE. For events occurring on the day of doses except for the first dose of study treatment, the events will be considered to occur prior to the study drug injection.

In general, summaries will include the number of patients in the safety population (*N*), frequency of patients experiencing the event (*n*), and relative frequency (that is, percentage; n/N*100).

In an AE overview table, the number and percentage of patients in the safety analysis set who experienced death, a serious adverse event (SAE), any TEAE, permanent discontinuation from study drug due to an AE, or a severe TEAE will be summarized by treatment group.

The number and percentage of patients with TEAEs will be summarized by treatment group in 2 formats listed below. For events that are gender specific, the denominator and computation of the percentage will only include patients from the given gender.

- By MedDRA PT nested within SOC with SOCs ordered alphabetically, and events ordered within each SOC by decreasing frequency in the 1000 mg of peresolimab Q4W treatment group
- By MedDRA PT with events ordered by decreasing frequency in the 1000 mg of peresolimab Q4W treatment group

AEs leading to permanent discontinuation of study drug and AEs leading to temporary interruption of study drug will also be summarized by treatment group using MedDRA PT nested within SOC. AE leading to temporary interruption of study drug will be identified using the response of “Drug Interrupted” in the question “What action taken with study treatment?” on the *Adverse Event* eCRF. Events will be ordered by decreasing frequency within SOC in 1000 mg of peresolimab Q4W treatment group.

A summary of temporary interruptions of study drug due to AEs will also be provided, showing the number of patients who experienced at least 1 temporary interruption and the number of temporary interruptions per patient with an interruption. Further, the duration of each temporary interruption (in days) and the cumulative duration of dose interruption (in days) using basic descriptive statistics (n, mean, SD, minimum, first quartile, median, third quartile, and maximum) will be displayed. Duration of study drug interruption will be calculated as: visit date of the missed dose – date of next resumed dose + 1, only if both dates are provided.

The number and percentage of patients with TEAEs will be summarized by maximum severity, SOC and PT by treatment using MedDRA PT ordered by decreasing frequency for the TEAEs. For each patient and TEAE, the maximum severity for the MedDRA level being displayed is the maximum postbaseline severity observed from all associated LLTs mapping to that MedDRA PT.

Adverse events of special interest include hypersensitivity, injection site reactions, infections, and malignancy.

4.6.2.1. Serious Adverse Events

An individual listing of all AEs including preexisting conditions will be provided. A separate listing will include AEs that led to permanent discontinuation from the study drug.

With the International Conference on Harmonisation (ICH) E2A guideline, a SAE is any AE that results in 1 of the following outcomes:

- death
- initial or prolonged inpatient hospitalization
- a life-threating experience (that is, immediate risk of dying)
- persistent or significant disability/incapacity
- congenital anomaly/birth defect
- considered significant by the investigator for any other reason

The number and percentage of patients who experienced any ICH-defined SAE will be summarized by treatment group during the treatment and follow-up periods using the MedDRA PT nested within SOC. The SAEs will be summarized by treatment group using MedDRA PT without SOC. An individual listing of all SAEs will be provided.

4.6.2.2. Allergic Reactions and Hypersensitivities

The current standard MedDRA Standardized Medical Queries (SMQs) to search for relevant events will be used. TEAEs are characterized as follows:

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

The frequency and percentage of patients experiencing treatment-emergent hypersensitivities will be summarized by treatment arm, SMQ, and PT during the Treatment Period. The frequency and percentage of patients experiencing treatment-emergent hypersensitivities will also be summarized by treatment arm and timing category.

Exposure-adjusted incidence rates (per 100 PYs) of treatment-emergent hypersensitivities will be summarized.

Total number of patients who experienced the TEAE for each PT while they were on Peresolimab will be divided by the sum of all patients' time (in 100 years) of Peresolimab

exposure during the treatment period. The entire time on the study during the Peresolimab treatment period will be used to calculate the patient's time of exposure.

The time to onset of first treatment-emergent hypersensitivity during the Treatment Period and the duration of each treatment-emergent hypersensitivity during the Treatment Period will also be defined and calculated in the same way.

A by-patient listing of all treatment-emergent hypersensitivities will be provided.

4.6.2.3. Injection Site Reactions

Injection site reactions will be defined using terms from the injection site reactions high level term. Injection site reactions will be captured in the eCRF as an AE.

The frequency and percentage of patients with a treatment-emergent injection site reaction will be summarized by treatment arm, maximum severity, SOC, and PT during Treatment Period.

The frequency and percentage of patients with injection site reactions will be summarized by treatment arm and by the following categories:

- Erythema
  - Severity (Noticeable but Very Mild Redness, Clearly Red, Bright Red, Dark with Ulceration or Necrosis at the Injection Site)
  - Size (Barely Noticeable, Slight, Moderate, Severe)
- Induration
  - Severity (Barely Noticeable, Slight, Moderate, Severe)
- Pain
  - Severity (Mild, Moderate, Severe)
- Pruritus
  - Severity (Mild, Moderate, Severe)
- Edema
  - Severity (Mild, Moderate, Severe)

A by-patient listing of treatment-emergent injection site reactions will be provided.

4.6.2.4. Infections

Infectious events include infections, serious infections, opportunistic infections, infections that require therapeutic intervention, and any events involving reactivation of tuberculosis or hepatitis. Such events will be selected using the 'Infections and Infestations' SOC. CCI [redacted]

Summary of Opportunistic Infections based on MedDRA PTs after the potential opportunistic infections are reviewed by medical and confirmed as opportunistic infections. Listing of Opportunistic Infections based on MedDRA PTs during the study.

CCI [redacted]

CCI

b. Localized herpes zoster infections are not considered OI; only multidermatomal and/or disseminated infections are considered OI.
    a. Localized or nonmultidermatomal are defined as involvement of the primary and/or adjacent dermatomes only. These may be complicated or uncomplicated:
        i. Complicated: documented ocular (cornea or deeper structure; for example, iritis, keratitis, retinitis, and so on) or motor nerve involvement (for example, palsy). Postherpetic neuralgia does not meet criteria for motor nerve involvement.
        ii. Uncomplicated: localized or nonmultidermatomal cases that are not complicated.
    b. Multidermatomal is defined as involvement beyond primary and adjacent dermatomes (that is, 4 or more contiguous dermatomes) or involvement of 2 or more noncontiguous dermatomes. These may be complicated or uncomplicated.
        i. Complicated: documented ocular (cornea or deeper structure; for example, iritis, keratitis, retinitis, and so on) or motor nerve development.
        ii. Uncomplicated: multidermatomal cases.
    c. Disseminated: systemic infection, visceral, or widespread cutaneous (for example, 5 or more dermatomes or from 3 to 4 dermatomes including at least 1 noncontiguous [nonadjacent]).
    d. Recurrent: Herpes zoster (HZ) infection is considered an opportunistic infection irrespective of number of dermatomal involvement when it recurs. Recurrent HZ is defined as either a reported HZ infection in the patient's medical history or prior to the study initiation, with at least one additional event occurring during the course of participation in the study. Alternatively, if there is no reported prior history of HZ, recurrent HZ is characterized by more than one HZ infection in an individual patient during the course of participation in the study.

CCI

The frequency and percentage of patients experiencing infectious AEs will also be summarized by treatment, maximum severity, SOC, and PT during the Treatment Period.

A listing of patients with detectable hepatitis B virus (HBV) DNA will be provided.

HBV DNA status (not detectable, detectable but not quantifiable [that is, <lower limit of detection (LLOD)], quantifiable [that is, ≥LLOD]) will be summarized, stratified by applicable baseline HBV serology status.

A summary table of herpes zoster will be provided, including event maximum severity, seriousness, whether resulting in temporary study drug interruption, whether resulting in study drug discontinuation, and event outcome. See Appendix 10 (Section 6.10) for PT list for herpes zoster.

A summary table of herpes simplex will be provided, including event maximum severity, seriousness, whether resulting in temporary study drug interruption, whether resulting in study drug discontinuation, and event outcome. See Appendix 10 (Section 6.10) for PT list for herpes simplex.

The time to onset of first treatment-emergent infectious AE during the Treatment Period will be defined as follows:

Time to onset of first treatment-emergent infectious AE (in weeks) =
(Start date of first infectious AE – Date of first injection +1) / 7.

Note that if an infectious AE was preexisting prior to the date of first injection and subsequently became treatment-emergent due to worsening post-baseline, the date of worsening will be used in place of the AE start date.

If a patient does not have a treatment-emergent infectious AE during the Treatment Period, they will be censored at the date of completion or ED from the Treatment Period. Kaplan-Meier estimates of the proportion of patients not yet experiencing their first treatment-emergent infectious AE at 12, 24, 36, 48, and 60 weeks will be provided for each treatment arm. A Kaplan-Meier plot of the time to onset will be presented by treatment arm.

The duration of each treatment-emergent infectious AE during the Treatment Period will be defined as follows:

Duration of treatment-emergent infectious AE (in weeks) =
(End date of AE – Start date of AE + 1) / 7.

If an AE has not ended by the date of completion or ED from the study, including the Posttreatment Follow-up Period, it will be censored as of that date. Descriptive statistics for the duration of treatment-emergent infectious AEs will be summarized by treatment arm. If a patient has more than 1 treatment-emergent infectious AE during the Treatment Period, then each AE will contribute to the analysis. If a patient does not have a treatment-emergent infectious AE during the Treatment Period, they will not be included in the analysis.

A by-patient listing of treatment-emergent infections will be provided.

4.6.2.5. Malignancies

Malignancies will be defined using terms from the Malignant tumors SMQ (20000194) or Malignancies SMQ (20000090).

The frequency and percentage of patients with a treatment-emergent malignancy will be summarized by treatment arm, maximum severity, SOC, and PT during the Treatment Period. Exposure-adjusted incidence rates (per 100 PYs) for each type of treatment-emergent malignancy will be presented by treatment arm, and PT during the Treatment Period. AEs will be summarized under a particular time interval if they started or worsened on or after the interval start date and before the interval end date.

Malignancies excluding nonmelanoma skin cancers (NMSC) and NMSC will be also reported separately. NMSC will be defined using the below PTs:

- Squamous cell carcinoma of skin (10041834)
- Bowen's disease (10006059)
- Basal cell carcinoma (10004146)
- Basosquamous carcinoma (10004178)
- Basosquamous carcinoma of skin (10004179)
- Squamous cell carcinoma (10041823)
- Skin squamous cell carcinoma metastatic (10077314)
- Skin cancer (10040808)
- Carcinoma in situ of skin (10007390)
- Keratoacanthoma (10023347)
- Vulvar squamous cell hyperplasia (10079905)
- Skin squamous cell carcinoma recurrent (10081136)
- Lip squamous cell carcinoma (10064055)
- Penile Squamous cell carcinoma (10059631)

The time to onset of first treatment-emergent malignancies during the Treatment Period will be summarized in the same way as time to onset of first treatment-emergent infections.

A by-patient listing of treatment-emergent malignancies will be provided.

Additional information will be provided in the Adjudication Committee Charter.

4.6.2.6. Depression

Depression will be defined using terms from the Depression and suicide/self-injury SMQ (20000035).

The frequency and percentage of patients with treatment-emergent depression will be summarized by treatment arm, maximum severity, SOC, and PT during the Treatment Period. Exposure-adjusted incidence rates (per 100 PYs) for each type of treatment-emergent depression will be presented by treatment arm, PT, and time interval during the Treatment Period. AEs will be summarized under a particular time interval if they started or worsened on or after the interval start date and before the interval end date.

A by-patient listing of treatment-emergent depression will be provided.

4.6.2.7. Suicide-related Thoughts and Behaviors

Suicide-related ideations and behaviors, based on Columbia Suicide Severity Rating Scale (C-SSRS), will be summarized by treatment arm separately for patients who completed the C-SSRS starting at enrollment versus patients who were already enrolled when the C-SSRS was introduced to the study during the Treatment Period.

Treatment-emergent suicide-related ideations and behaviors during the Treatment Period are defined as any ideation or behavior reported after the baseline visit (Visit 2) and on or prior to Visit 18 (Week 60) or ED. The frequency and percentage of patients with treatment-emergent

suicide-related ideations and behaviors will be summarized by treatment arm for patients who completed the C-SSRS starting at enrollment by C-SSRS category (1 through 10), by type (ideation or behavior), and overall during the Treatment Period.

Suicide-related ideations and behaviors will be listed by patient and visit separately for patients who completed the C-SSRS starting at enrollment versus patients who were already enrolled when the C-SSRS was introduced to the study. Only patients that show suicidal ideation/behavior will be displayed, including lifetime history at screening. If patient's answers are all 'no' for the C-SSRS, then that patient will not be displayed. However, if a patient reported any ideation or behavior at any time point, then all their ideation and behavior will be displayed, even if not positive.

4.6.3. Clinical Laboratory Evaluation

All numerical laboratory tests for serum chemistries, fasting lipids, hematology, rheumatoid factor (RF) and anti-CCP antibodies, CRP, and ESR will be summarized by treatment arm for observed and change from baseline values during the Treatment Period.

Urinalysis will appear in data listings.

Summary of B-cells, immunoglobulins (Igs), and immunogenicity tests are described in the subsections below.

Change from baseline in laboratory tests will be summarized for patients who have both a baseline and at least postbaseline result.

Change from the minimum value on or prior to the date of the first injection of study treatment to the minimum value during the Treatment Period in laboratory tests will be summarized for patients who have both a baseline and at least 1 postbaseline result. Similarly, change from the maximum value on or prior to the date of the first injection of study treatment to the maximum value during the Treatment Period in laboratory tests will be summarized for patients who have both a baseline and at least 1 post-baseline result. Scheduled visits, unscheduled visits, and repeat measurements will be included.

The frequency and percentage of patients with treatment-emergent abnormal, high, or low laboratory results at any time will be summarized. Similarly, the frequency and percentage of patients with treatment-emergent abnormal, high, or low laboratory results will be summarized. A treatment-emergent abnormal result is defined as a change from normal at all baseline visits to abnormal at any time during the Treatment Period. A treatment-emergent high result is defined as a change from a value less than or equal to the high limit at all baseline visits to a value greater than the high limit at any time during the Treatment Period. A treatment-emergent low result is defined as a change from a value greater than or equal to the low limit at all baseline visits to a value less than the low limit at any time during the Treatment Period. For each laboratory test, only patients who were normal (that is, less than or equal to the high limit for treatment-emergent high, or greater than or equal to the low limit for treatment-emergent low) at all baseline visits and who have at least 1 nonmissing postbaseline result will be included in the denominator when computing the percentages of patients with treatment-emergent abnormal,

high, or low results. Scheduled visits, unscheduled visits, and repeat measurements will be included. Covance Reference Ranges will be used to define the low and high limits. Listing of significant lab test abnormalities will be provided.

Shifts from baseline to maximum postbaseline result during the Treatment Period for alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP) and total bilirubin will be provided for the following categories:

- ALT, AST; ≤1 × upper limit of normal (ULN), >1 to <3 × ULN, ≥3 to <5 × ULN, ≥5 to <10 × ULN, and ≥10 × ULN
- ALP; ≤1 × ULN, >1 to <1.5 × ULN, ≥1.5 to <2 × ULN, ≥2 × ULN
- Total bilirubin; ≤1 × ULN, >1 to <1.5 × ULN, ≥1.5 to <2 × ULN, ≥2 × ULN

Lab Evaluation of Drug-Induced Serious Hepatotoxicity plot will be created with maximum postbaseline ALT/AST divided by ULN vs. maximum postbaseline total bilirubin divided by ULN.

Hematologic changes will be defined based on clinical laboratory assessments. Common Terminology Criteria for Adverse Events (CTCAEs) will be applied for selected laboratory tests. Shifts from maximum baseline to maximum postbaseline result will be provided by treatment periods for the categories in Appendix 11 (Section 6.11).

Shift tables from maximum baseline to maximum postbaseline for lipid effects that include total cholesterol and triglycerides will be categorized based on NCEP levels in Appendix 11 (Section 6.11) and summarized by treatment periods.

Shift table from maximum baseline to maximum postbaseline for renal function that includes elevated creatinine will be categorized based on CTCAE in Appendix 11 (Section 6.11) and summarized by treatment periods.

Scheduled visits, unscheduled visits, and repeat measurements will be included.

4.6.4. Pharmacodynamics

4.6.4.1. Flow Cytometry

For flow cytometry panel, the absolute count and relative count (that is, as a percentage of the total lymphocyte population) will be summarized by treatment arm for baseline and all postbaseline visits during the Treatment Period using descriptive statistics.

The change from baseline and the percent change from baseline to each postbaseline visit during the Treatment Period will be summarized for absolute and relative count. Note that the baseline count is defined as the average of the nonmissing results on or prior to the date of the first injection of blinded study treatment, unless 1 is missing, in which case the nonmissing result will be used. Unscheduled and repeated measurements will be included.

In addition, the mean absolute and relative counts during the Treatment Period will be presented graphically over time (in weeks) by treatment arm. Similarly, the mean percentage change from

baseline in absolute and relative counts will be presented graphically over time (in weeks) by treatment arm.

A by-patient listing of cytometry cell counts during the Treatment Period will be provided.

4.6.4.2. Serum Immunoglobulin Levels and Cytokines

Serum Ig levels, including IgA, IgG, and IgM, and cytokines will be summarized by treatment arm for baseline, postbaseline visits, and both change from baseline and percent change from baseline to each postbaseline visit. Baseline is defined as the last nonmissing result on or prior to the date of the first injection of study treatment.

For each parameter, the mean percentage change from baseline over time (in weeks) will be summarized graphically by treatment arm.

Shift tables will be produced showing the frequency and percentage of patients with a minimum postbaseline result during the Treatment Period in each relevant category, by treatment arm and baseline result. Such shift tables will be produced for the following parameters and categories:

- IgA; <LLN to ≥ 0.35 g/L, and <0.35 g/L
- IgG: <LLN to ≥2.82 g/L, and <2.82 g/L
- IgM: <LLN to ≥0.20 g/L, and <0.20 g/L

A by-patient listing of serum Ig levels and cytokines during the Treatment Period will be provided.

CCI [redacted]

4.6.4.4. Immunogenicity

A participant is defined as evaluable for TE ADA if the participant has a nonmissing baseline ADA result and at least 1 nonmissing postbaseline ADA result.

A participant who is evaluable for TE ADA is defined as TE ADA+ if either of the following holds:

CCI [redacted]

A NAb+ participant is defined as a TE ADA+ participant where an NAb Present result is detected for at least 1 postbaseline ADA sample.

The frequency and percentage of TE ADA+ and NAb+ patients will be summarized by treatment arm.

The time to the development of TE ADA will be calculated as:

Time to development of TE ADA (in days) = Date of development of TE ADA – Date of first injection + 1.

If a patient has not developed TE ADA, they will be censored at the date of the last immunogenicity assessment during the Treatment Period.

The frequency and percentage of patients with both a TE ADA+ incidence (yes, no) and at least 1 TE hypersensitivity AE (yes, no) will be summarized by treatment arm. Similarly, the frequency and percentage of patients with both a TE ADA+ incidence (yes, no) and at least 1 treatment-emergent injection site reaction (yes, no) AE will be summarized by treatment arm.

4.6.5. Vital Signs and Other Physical Findings

Vital signs and physical characteristics to be summarized include temperature, sitting heart rate, sitting blood pressure (systolic and diastolic), weight, and body mass index (BMI). Vital signs and physical characteristics will be summarized by treatment arm for baseline, postbaseline visits, and change from baseline to postbaseline visits.

Change from baseline in vital signs and physical characteristics will be summarized for patients who have both a baseline and at least 1 postbaseline result. Baseline is defined as the last nonmissing result on or prior to the date of the first injection of study treatment for each period.

Change from the minimum value on or prior to the date of the first injection of study treatment to the minimum value during the Treatment Period in vital signs and physical characteristics will be summarized for patients who have both a baseline and at least 1 postbaseline result. Similarly, change from the maximum value on or prior to the date of the first injection of study treatment to the maximum value during the Treatment Period in vital signs and physical characteristics will be summarized for patients who have both a baseline and at least 1 postbaseline result. Scheduled visits, unscheduled visits, and repeat measurements will be included.

The percentages of patients with treatment-emergent high or low vital signs and physical characteristics at any time during the Treatment Period will be summarized. Similarly, the percentages of patients with treatment-emergent high or low vital signs and physical characteristics will be summarized. A treatment-emergent high result is defined as a change from a value less than the high limit at all baseline visits to a value greater than or equal to the high limit at any time during the Treatment Period. A treatment-emergent low result is defined as a change from a value greater than the low limit at all baseline visits to a value less than or equal to the low limit at any time during the Treatment Period. For each vital sign and physical characteristic, only patients who were normal (that is, less than the high limit for treatment-emergent high or greater than the low limit for treatment-emergent low) at all baseline visits and who have at least 1 nonmissing postbaseline result will be included in the denominator when computing the percentages of patients with treatment-emergent high or low results. Scheduled

visits, unscheduled visits, and repeat measurements will be included. Below table will be used to define the low and high limits.

Parameter	Low	High
Systolic BP (mm Hg)[a] (supine or sitting – forearm at heart level)	≤90 and decrease from baseline ≥20	≥160 and increase from baseline ≥20
Diastolic BP (mm Hg)[a] (supine or sitting – forearm at heart level)	≤50 and decrease from baseline ≥10	≥100 and increase from baseline ≥10
Heart rate (bpm)[a] (supine or sitting)	<50 and decrease from baseline ≥15	>120 and increase from baseline ≥15
Weight (kg) (consistent clothing and timing in relationship to meals and voiding)	(Loss) decrease from baseline ≥7%	(Gain) increase from baseline ≥7%

Abbreviation: BP = blood pressure.

A by-patient listing of all vital signs and physical characteristics during the Treatment Period will be provided. Additionally, a by-patient listing of treatment-emergent high or low vital signs and physical characteristics during the Treatment Period will be provided.

4.6.6. Pregnancy

Any reported positive pregnancy results or reported pregnancies will be listed.

4.7. Other Analyses

4.7.1. Efficacy Subgroup Analyses

Subgroup analyses of the primary endpoint will be made to assess consistency of the intervention effect across the following subgroups (but not limited to only these):

- Prior b/tsDMARD failure (Yes vs. No; number of failures: 0, 1, >1)
- Prior b/tsDMARD (ever vs. never received)
- Gender
- Race
- Geographic region
- Concomitant medication use
- RF and/or anticitrullinated peptide antibodies positive, and
- Disease duration (i.e. <6 months, 6 months to 2 years, >2 to 5 years, >5 to 10 years, >10 years).

Descriptive statistics will be provided for each treatment and stratum of a subgroup as outlined, regardless sample size.

Categorical variable will be analyzed with a logistic regression model that contains the treatment, the subgroup variable, and subgroup by treatment interaction. The treatment-by-subgroup

interaction will be tested at the 10% significant level. Within each subgroup category, the proportion of responders by treatment, treatment differences, and 95% CIs will be displayed.

If the number of patients in any subgroup category is less than 10% of the total population, only summaries of the efficacy data will be provided. Additional subgroup analyses may be performed as deemed necessary.

4.7.2. Safety Subgroup Analyses

Safety subgroup analysis for safety related endpoints will be performed within the context of the integrated safety analysis. No safety subgroup analysis will be performed specifically for this study unless there is a potentially relevant finding during the periodic study safety reviews.

4.7.3. Pharmacokinetic/Pharmacodynamic Analyses

Peresolimab concentrations will be illustrated graphically and summarized descriptively.

A model-based approach using nonlinear mixed effects modeling or other appropriate software may be conducted. If appropriate, data from this study may be combined with data from other studies in model-based analyses. The impact of extrinsic and intrinsic patient factors, such as age, weight, and sex, on PK may be examined as needed. Estimates of PK model parameters and covariate effects will be reported. If ADA titers are detected from immunogenicity testing, then the impact of immunogenicity titers on peresolimab PK parameters may also be examined.

If data warrant, analyses of exposure-response relationships may be conducted using both exploratory graphical approaches and model-based approaches.

Exploratory graphical analysis approaches for categorical clinical endpoints may consist of graphs showing the percentage of participants that achieve the clinical endpoint at different percentiles, for example, quartiles, of exposure of peresolimab at Week 12. Measures of exposure may include population PK estimated average concentrations at steady state ($C_{avg, ss}$) or observed trough concentrations at the time of the clinical endpoint.

Model-based analyses of the categorical clinical endpoints may use population exposure-response logistic regression models, where maximum effect or other model structures may be used to relate exposure to the probability of achieving the endpoint. These models may be used to evaluate participant factors that may impact the relationship between exposure and the probability of achieving the endpoint. Longitudinal exposure-response models for ACR scores or response rates may be developed, which relate the time course and magnitude of peresolimab exposure to the time course and magnitude of ACR response.

Exploratory PK/PD analyses may be conducted to evaluate the relationship between peresolimab exposure and select measures of response, such as RO. Additional analyses may be conducted if they are deemed appropriate.

4.7.4. Analysis for Japan Submission

A subset of the planned efficacy, health outcomes, and safety analyses will be reproduced based on patients from Japan sites (Japanese population), in support of the regulatory submission in

Japan. The list of tables, figures, and listings for the Japanese population will be provided in a separate document.

4.8. Interim Analyses

4.8.1. Data Monitoring Committee or Other Review Board

There will be no Data Monitoring Committee in this study. An assessment of unblinded interim data will be conducted by an internal assessment committee with a limited number of prespecified team members who do not have direct site contact or data entry or validation responsibilities. An assessment committee (AC) charter will provide details on AC membership and the governing processes.

4.9. Changes to Protocol-Planned Analyses

There is no change to protocol-planned analyses.

5. Sample Size Determination

Approximately 420 participants CCI

- 100 mg of peresolimab Q4W
- 400 mg of peresolimab Q4W
- 1000 mg of peresolimab Q4W
- Placebo to 400 mg of peresolimab Q4W, or
- Placebo to 1000 mg of peresolimab Q4W.

CCI

- At Week 12, the 100 mg peresolimab treatment group will have a sample size of approximately 60 participants, and the 400 mg and 1000 mg peresolimab treatment groups will each have a sample size of approximately 120 participants.
- At Week 12, the placebo group will have a sample size of approximately 120 participants.

CCI

CCI

6. Supporting Documentation

6.1. Appendix 1: Demographic and Baseline Characteristics

The patient's year of birth, sex, weight, height, previous b/tsDMARD treatment, and other demographic characteristics are collected at the screening visit. Age and BMI will be calculated.

Only the year of birth is collected at screening. For the purpose of age calculation, the month and day of birth will be imputed as July 01, of the year of birth. Age is computed as follows;

Age = (Informed Consent Date –Date of Birth +1)/365.25.

Demographic and baseline characteristics (including age, gender, race, and ethnicity) will be summarized for each treatment group.

Certain characteristics that are collected at baseline or after baseline but not summarized in the demographic summary will be reported as a listing.

No inferential analysis for the comparability of baseline covariates across treatment groups will be performed.

6.2. Appendix 2: Treatment Compliance

CCI [redacted]

Study treatment administration and compliance will be listed for all entered patients. The number and percentage of patients who are treatment compliant by week (that is, at each injection time point) will be summarized by treatment group for each treatment period.

No patient will be excluded from the mITT population as a consequence of significant noncompliance.

Participants who are noncompliant with treatment will be listed by treatment group.

No analyses are planned to assess treatment compliance.

6.3. Appendix 3: Clinical Trial Registry Analyses

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

Analyses provided for the CTR requirements include the following:

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

- An AE is considered 'Serious' whether or not it is a TEAE.
- An AE is considered in the 'Other' category if it is both a TEAE and is not serious. For each SAE and 'Other' AE, for each term and treatment group, the following are provided:
    - the number of participants at risk of an event

Approved on 01 Dec 2024 GMT

    - the number of participants who experienced each event term
    - the number of events experienced.
- Consistent with www.ClinicalTrials.gov requirements, 'Other' AEs that occur in fewer than 5% of patients/subjects in every treatment group may not be included if a 5% threshold is chosen (5% is the minimum threshold).

AE reporting is consistent with other document disclosures for example, the CSR, manuscripts, and so forth.

6.4. Appendix 4: Preexisting Conditions and Medical History

Preexisting condition is defined the condition/event recorded on the Preexisting Conditions and Medical History eCRF page with a start date prior to the date of informed consent and no end date (that is, the event is ongoing) or an end date on or after the date of informed consent. In addition, the AEs occurring prior to first dose are also included. Notice if a preexisting condition worsens in severity on or after the date of informed consent, it will be recorded as an AE on the AE eCRF page with the date of worsening as the start date. The number and percentage of patients with preexisting conditions will be summarized by treatment group using the MedDRA PT nested within SOC. Summaries will be performed for the Safety population.

If year of start and end date of medical history events is missing, start date and end date will be considered as missing. If month of start and end date is missing, missing month will be taken as January and December, respectively. If date of start and end date is missing, missing date will be taken as the first and last day of the month, respectively.

6.5. Appendix 5: Prior and Concomitant Medication

Prior and concomitant medications will be summarized by treatment group and will be classified into anatomical therapeutic chemical drug classes using the latest version of the World Health Organization drug dictionary. Medication start and stop dates will be compared to the date of first dose of treatment in each treatment period to allow medications to be classified as Concomitant for each treatment period.

Prior medications are those medications that start and stop prior to the date of first dose of study treatment. Concomitant medications are those medications that start before, on, or after the first day of study treatment of the defined treatment period and continue into the treatment period.

If year of start and stop date is missing, it will not be considered as prior and concomitant medications. If month of start and stop date is missing, missing month will be taken as January and December, respectively. If date of start and stop date is missing, missing date will be taken as the first and last day of the month, respectively.

6.6. Appendix 6: Algorithm for Determining American College of Rheumatology Response

Details presented in this appendix will use "x" as a generic symbol, and the appropriate number (either 20, 50, or 70) is to be filled in when implementing in dataset programming code.

ACRx response is defined as $\geq x\%$ improvement from baseline in TJC (68 joint count), and $\geq x\%$ improvement in SJC (66 joint count), and $\geq x\%$ improvement in at least 3 of the following items:

- Patient's Assessment of Arthritis Pain
- Patient's Global Assessment of Disease Activity
- Physician's Global Assessment of Disease Activity
- Patient's Assessment of Physical Function as measured by the HAQ-DI, and
- Acute Phase reactant as measured by hsCRP.

The following abbreviations will be used throughout this appendix to refer to the items needed in the algorithm definitions:

Parameter	**Abbreviation for the Parameter**
% improvement in tender joint count	TJC68
% improvement in swollen joint count	SJC66
% improvement in patient's assessment of arthritis pain	Pain
% improvement in patient's global assessment of disease activity	PatGA
% improvement in physician's global assessment of disease activity	PGA
% improvement in HAQ-DI	HAQDI
% improvement in hsCRP	hsCRP

Abbreviations: HAQ-DI = Health Assessment Questionnaire - Disability Index; hsCRP = high-sensitivity C-reactive protein.

For all 7 parameters mentioned above, % improvement at a visit is calculated as:

$$\frac{\text{baseline value} - \text{value at visit}}{\text{baseline value}} \times 100.$$

If a baseline value or the value at the visit is missing for a particular parameter, then the percentage improvement is defined as missing.

To calculate the observed ACRx response at a visit

1. If the patient discontinued from the study prior to reaching the visit, then STOP and assign ACRx response as missing. Otherwise, calculate the percentage improvement at the visit for all 7 parameters as described above.
2. Consider the following joint count parameters: TJC68 and SJC66.
   - If TJC68 and SJC66 are BOTH $\geq x\%$, then proceed to Step 3.
   - If both TJC68 and SJC66 are nonmissing, but 1 or both is $< x\%$, then STOP and assign the patient as a nonresponder for ACRx.
   - If one of TJC68 or SJC66 is nonmissing, and the non-missing value is $< x\%$, then STOP and assign the patient as a nonresponder for ACRx.
   - If one of TJC68 or SJC66 is nonmissing, and the nonmissing value is $\geq x\%$, then STOP and assign ACRx response as missing.
   - If both are missing, then STOP and assign ACRx response as missing.

3. Consider the following 5 parameters: Pain, PatGA, PGA, HAQDI, and hsCRP.
   - If 3 or more items are missing, then STOP and assign ACRx response as missing.
   - If 3 or more items are nonmissing AND if at least 3 items are $\geq x\%$, then STOP and assign the patient as a responder for ACRx; otherwise, assign the patient as a nonresponder for ACRx.

6.7. Appendix 7: Details of Joint Count, HAQ-DI, and FACIT-F

Joint assessment

Each of 28 or 68 joints will be evaluated for tenderness and 28 or 66 joints will be evaluated for swelling at the specified visits as shown in the SoA in the protocol.

The 68 joints to be assessed and classified as tender or not tender include:

- 2 temporomandibular joints
- 2 sternoclavicular joints
- 2 acromioclavicular joints
- 2 shoulder joints
- 2 elbow joints
- 2 wrist joints
- 2 hip joints
- 2 knee joints
- 2 ankle joints
- 2 tarsal joints
- Hands:
  - 10 metacarpophalangeal joints
  - 2 interphalangeal joints of the thumb
  - 8 proximal interphalangeal joints
  - 8 distal interphalangeal joints
- Feet:
  - 10 metatarsophalangeal joints
  - 2 first proximal interphalangeal joints
  - 8 proximal interphalangeal joints

The 66 joints to be assessed and classified as swollen or not swollen are the same as those for the TJC except for the 2 hip joints.

For scores using 28 joints, the following subset will be assessed:

- 2 shoulder joints
- 2 elbow joints
- 2 wrist joints
- 2 knee joints
- 10 metacarpophalangeal joints
- 2 interphalangeal joints of the thumb

- 8 proximal interphalangeal joints

Calculating the joint counts

- A patient must have at least 50% of joints evaluated for swelling or tenderness. If more than 50% joints are "not evaluable" or "not done", then a joint count will not be computed.
- If all joints are evaluated, then the joint count score will be the number of joints for swelling or tenderness.
- If a patient has either "not evaluable" or "not done" joints, then the joint count score will be prorated and rounded up. For example, the TJC68 score for 15 tenderness with 3 "not evaluable" and 2 "not done" is [15/(68-3-2)]*68 = 16.19. Therefore, the TJC68 score for this patient will be 17.

HAQ-DI

The HAQ-DI is a patient-reported questionnaire that is commonly used in RA to measure disease-associated disability (assessment of physical function). It consists of 20 questions referring to 8 domains: dressing/grooming, arising, eating, walking, hygiene, reach, grip, and activities.

- Dressing/grooming includes 2 component questions with 1 aid/device checkbox and 1 help checkbox.
- Arising includes 2 component questions with 1 aid/device checkbox and 1 help checkbox.
- Eating includes 3 component questions with 1 aid/device checkbox and 1 help checkbox.
- Walking includes 2 component questions with 4 aid/device checkboxes and 1 help checkbox.
- Hygiene includes 3 component questions with 4 aid/device checkboxes and 1 help checkbox.
- Reach includes 2 component questions with 1 aid/device checkbox and 1 help checkbox.
- Grip includes 3 component questions with 1 aid/device checkbox and 1 help checkbox.
- Activity includes 3 component questions with 1 help checkbox.

Participants assess their degree of difficulty when performing the above activities over the past week. The following scores are assigned to the responses:

- Without any difficulty = 0
- With some difficulty = 1
- With much difficulty = 2
- Unable to do = 3

The reported use of special aids or devices and/or the need for assistance of another person to perform these activities is also assessed.

Calculating the HAQ-DI

The patient must have a score for at least 6 of the 8 categories. If there are less than 6 categories completed, a HAQ-DI cannot be computed.

- A category score is determined from the highest score of the component questions in that category. For instance, in the category of Hygiene, there are 3 component questions. If a patient responds with a 0, 1, 3, respectively, then the category score is 3.
- If there is no aid/device and help needed for a category, then the category score is not modified.
- If there is at least 1 aid/device or help needed for a category, then the category score will be modified.
    - When patient's category score is either 0 or 1, the category score will be adjusted to 2.
    - When patient's category score is either 2 or 3, the category score will remain 2 or 3, respectively, as its original category score.
- Average the evaluable category scores.

The HAQ-DI functional disability index score ranges from 0 to 3 by increment 0.125.

FACIT-F score

The FACIT-F score (Cella and Webster 1997) is a brief, 13-item, symptom-specific questionnaire that specifically assesses the self-reported severity of fatigue and its impact upon daily activities and functioning. The FACIT-F questionnaire is presented below.

A 5-point Likert-type scale (0 = not at all; 1 = a little bit; 2 = somewhat; 3 = quite a bit; 4 = very much) is used. As each of the 13 items of the FACIT-F scale ranges from 0 to 4, the range of possible score is 0 to 52, with 0 being the worst possible score and 52 being the best.

Calculating the FACIT-F score

The participant must have a score for more than 7 items. If less than 7 items are answered, the FACIT-F score will be considered missing.

- Reverse scores for 11 out of the 13 items except for the 2 questions, "I have energy" and "I am able to do my usual activities."
- Sum the individual items to obtain a score.
- Multiply the sum by 13, then divide by the number of evaluable items.

6.8. Appendix 8: List of Geographic Region

Geographic region will be used as one of stratification factors for randomization at Week 0. Patients will be classified as one of the following geographic regions based on the sites where patients are enrolled: Europe, North America, and Other.

The list of countries in each geographic region are:

- Europe: Greece, Hungary, Poland, Spain

- North America: United States, Canada, and
- Other: Argentina, China, Japan, Mexico.

6.9. Appendix 9: Execution for Bayesian Model Averaging

6.9.1. R DREAMER Package to Execute BMA

BMA will be executed using the R package *DREAMER*: **D**ose **RE**sponse b**A**yesian **M**odel averaging. *DREAMER* package and supporting documentation may be found at GitHub - rich-payne/dreamer and on CRAN.

6.9.2. BMA Components, Priors, and Weights

E-max, log-linear, log-quadratic, and exponential dose response models are the prespecified component to the BMA model that will be used to analyze the KDAF dose response and are described in the DREAMER supporting documentation. BMA components and weights are summarized in below table. The priors for each parameter are normal distributions, except b_4 in the Sigmoidal EMAX model, which is truncated normal (to be positive). A logistic link function will be used for each model.

Component	Weight	Model	Parameter	mu	sigma
Sigmoidal EMAX	1/4	$f(d) = b_1 + \frac{(b_2 - b_1)d^{b_4}}{\exp(b_3 b_4) + d^{b_4}}$	b_1	-2	1.2
			b_2	2	1.25
			b_3	log(500)	2
			b_4	0	5
Log-Linear	1/4	$f(d) = b_1 + b_2 \log(d + 1)$	b_1	-2	1.2
			b_2	0	0.5
Log-Quadratic	1/4	$f(d) = b_1 + b_2 \log(d + 1) + b_3 \log(d + 1)^2$	b_1	-2	1.2
			b_2	2	1
			b_3	0	1
Exponential	1/4	$f(d) = b_1 + b_2(1 - \exp(-b_3 d))$	b_1	-2	1.2
			b_2	2	1.75
			b_3	0	1

6.10. Appendix 10: List of Preferred Term for Herpes Zoster, Herpes Simplex, and Hepatitis B Virus

Category	Preferred Term (MedDRA Version 22.1)	Preferred Term Code	Lilly Defined Classification
Herpes Simplex	Colitis herpes	10051782	Narrow
	Eczema herpeticum	10014197	Broad
	Elsberg syndrome	10086845	Narrow
	Gastritis herpes	10051784	Narrow
	Genital herpes	10018150	Broad
	Genital herpes simplex	10073931	Broad
	Herpes dermatitis	10062639	Broad
	Herpes oesophagitis	10052330	Narrow
	Herpes ophthalmic	10062004	Narrow
	Herpes pharyngitis	10066888	Broad
	Herpes sepsis	10058876	Narrow
	Herpes simplex	10019948	Broad
	Herpes simplex bronchitis	10085017	Narrow
	Herpes simplex cervicitis	10077449	Narrow
	Herpes simplex colitis	10074239	Narrow
	Herpes simplex encephalitis	10019953	Narrow
	Herpes simplex gastritis	10074240	Narrow
	Herpes simplex hepatitis	10067389	Narrow
	Herpes simplex meningitis	10019956	Narrow
	Herpes simplex meningoencephalitis	10074247	Narrow
	Herpes simplex meningomyelitis	10074250	Narrow
	Herpes simplex necrotising retinopathy	10074252	Narrow
	Herpes simplex oesophagitis	10074242	Narrow
	Herpes simplex otitis externa	10019959	Narrow
	Herpes simplex pharyngitis	10074244	Broad
	Herpes simplex pneumonia	10065046	Narrow
	Herpes simplex sepsis	10074246	Narrow
	Herpes simplex test positive	10077969	Broad
	Herpes simplex viraemia	10080365	Narrow
	Herpes simplex virus conjunctivitis neonatal	10049458	Broad
	Herpes simplex virus urethritis	10087751	Narrow
	Herpes simplex visceral	10019963	Narrow

Category	Preferred Term (MedDRA Version 22.1)	Preferred Term Code	Lilly Defined Classification
	Herpes virus infection	10019973	Broad
	Lower respiratory tract herpes infection	10077390	Narrow
	Meningitis herpes	10027242	Narrow
	Meningoencephalitis herpes simplex neonatal	10053586	Broad
	Meningoencephalitis herpetic	10027285	Narrow
	Meningomyelitis herpes	10074249	Narrow
	Nasal herpes	10074936	Broad
	Neonatal mucocutaneous herpes simplex	10053587	Broad
	Ophthalmic herpes simplex	10073938	Narrow
	Oral herpes	10067152	Broad
	Pneumonia herpes viral	10035703	Narrow
	Proctitis herpes	10036780	Narrow
Herpes Zoster	Disseminated varicella zoster virus infection	10084396	Narrow
	Genital herpes zoster	10072210	Broad
	Herpes zoster	10019974	Broad
	Herpes zoster cutaneous disseminated	10074297	Narrow
	Herpes zoster disseminated	10065038	Narrow
	Herpes zoster infection neurological	10061208	Narrow
	Herpes zoster meningitis	10074259	Narrow
	Herpes zoster meningoencephalitis	10074248	Narrow
	Herpes zoster meningomyelitis	10074251	Narrow
	Herpes zoster meningoradiculitis	10079327	Narrow
	Herpes zoster necrotising retinopathy	10074253	Narrow
	Herpes zoster oticus	10063491	Narrow
	Herpes zoster pharyngitis	10074245	Narrow
	Herpes zoster reactivation	10080516	Narrow
	Herpetic radiculopathy	10082717	Narrow
	Necrotising herpetic retinopathy	10065119	Narrow
	Ophthalmic herpes zoster	10030865	Narrow
	Oral herpes zoster	10086594	Narrow
	Varicella zoster gastritis	10074241	Broad
	Varicella zoster oesophagitis	10074243	Broad
	Varicella zoster pneumonia	10074254	Broad
	Varicella zoster sepsis	10074298	Broad
	Varicella zoster viraemia	10087746	Broad

Category	Preferred Term (MedDRA Version 22.1)	Preferred Term Code	Lilly Defined Classification
	Varicella zoster virus infection	10075611	Broad
Tuberculosis	Adrenal gland tuberculosis	10001358	Narrow
	Bone tuberculosis	10056377	Narrow
	Bovine tuberculosis	10006049	Narrow
	Cardiac tuberculosis	10087547	Narrow
	Choroid tubercles	10008779	Narrow
	Conjunctivitis tuberculous	10010754	Narrow
	Cutaneous tuberculosis	10011684	Narrow
	Disseminated tuberculosis	10013453	Narrow
	Ear tuberculosis	10014027	Narrow
	Epididymitis tuberculous	10015004	Narrow
	Extrapulmonary tuberculosis	10064445	Narrow
	Female genital tract tuberculosis	10061150	Narrow
	Immune reconstitution inflammatory syndrome associated tuberculosis	10072797	Narrow
	Intestinal tuberculosis	10075268	Narrow
	Joint tuberculosis	10056367	Narrow
	Latent tuberculosis	10065048	Narrow
	Lymph node tuberculosis	10025183	Narrow
	Male genital tract tuberculosis	10061234	Narrow
	Mammary tuberculosis	10083169	Narrow
	Meningitis tuberculous	10027259	Narrow
	Mycobacterium tuberculosis complex test positive	10070325	Broad
	Oesophageal tuberculosis	10030200	Narrow
	Oral tuberculosis	10076879	Narrow
	Pericarditis tuberculous	10055069	Narrow
	Peritoneal tuberculosis	10053583	Narrow
	Prostatitis tuberculous	10064743	Narrow
	Pulmonary tuberculoma	10066927	Narrow
	Pulmonary tuberculosis	10037440	Narrow
	Renal tuberculosis	10038534	Narrow
	Salpingitis tuberculous	10039463	Narrow
	Silicotuberculosis	10068876	Broad
	Spleen tuberculosis	10041640	Narrow
	Thyroid tuberculosis	10043774	Narrow

Category	Preferred Term (MedDRA Version 22.1)	Preferred Term Code	Lilly Defined Classification
	Tuberculoma	10089070	Narrow
	Tuberculoma of central nervous system	10052883	Narrow
	Tuberculosis	10044755	Narrow
	Tuberculosis bladder	10044758	Narrow
	Tuberculosis gastrointestinal	10061390	Narrow
	Tuberculosis liver	10058120	Narrow
	Tuberculosis of central nervous system	10061391	Narrow
	Tuberculosis of eye	10044819	Narrow
	Tuberculosis of genitourinary system	10044828	Narrow
	Tuberculosis of intrathoracic lymph nodes	10044846	Narrow
	Tuberculosis of peripheral lymph nodes	10044965	Narrow
	Tuberculosis of uterine cervix	10086673	Narrow
	Tuberculosis ureter	10045026	Narrow
	Tuberculous abscess central nervous system	10052884	Narrow
	Tuberculous endometritis	10071559	Narrow
	Tuberculous laryngitis	10045072	Narrow
	Tuberculous pelvic inflammatory disease	10086676	Narrow
	Tuberculous pleurisy	10045104	Narrow
	Tuberculous tenosynovitis	10059161	Narrow

6.11. Appendix 11: Common Terminology Criteria for Adverse Events (CTCAE)

Selected Hematologic Events

Event	Laboratory Test	Grade	Criteria in Système International (SI) Units	Criteria in Conventional (CN) Units
Anemia*	Hemoglobin	0 (normal)	≥7.27 mmol (Fe)/L for females and ≥8.18 mmol (Fe)/L for males	≥12 g/dL for females and ≥13.5 g/dL for males
		1	<7.27 mmol (Fe)/L for females and 8.18 mmol (Fe)/L for males and ≥6.2 mmol (Fe)/L	<12 g/dL for females and 13.5 g/dL for males and ≥10 g/dL
		2	<6.2 mmol (Fe)/L and ≥4.9 mmol (Fe)/L	<10 g/dL and ≥8.0 g/dL
		3	<4.9 mmol (Fe)/L and ≥4.0 mmol (Fe)/L	<8.0 g/dL and ≥6.5 g/dL
		4	<4.0 mmol (Fe)/L	<6.5 g/dL
Leukopenia*	White blood cell (WBC) count	0 (normal)	≥4.0 billion cells/L	≥4.0 thousand cells/uL
		1	<4.0 billion cells/L and ≥3.0 billion cells/L	<4.0 thousand cells/uL and ≥3.0 thousand cells/uL
		2	<3.0 billion cells/L and ≥2.0 billion cells/L	<3.0 thousand cells/uL and ≥2.0 thousand cells/uL
		3	<2.0 billion cells/L and ≥1.0 billion cells/L	<2.0 thousand cells/uL and ≥1.0 thousand cells/uL
		4	<1.0 billion cells/L	<1.0 thousand cells/uL
Neutropenia*	Absolute neutrophil count (ANC)	0 (normal)	≥2 billion cells/L	≥2 thousand cells/uL
		1	<2 billion cells/L and ≥1.5 billion cells/L	<2 thousand cells/uL and ≥1.5 thousand cells/uL
		2	<1.5 billion cells/L and ≥1.0 billion cells/L	<1.5 thousand cells/uL and ≥1.0 thousand cells/uL
		3	<1.0 billion cells/L and ≥0.5 billion cells/L	<1.0 thousand cells/uL and ≥0.5 thousand cells/uL
		4	<0.5 billion cells/L	<0.5 thousand cells/uL
Lymphopenia*	Lymphocyte count	0 (normal)	≥1.1 billion cells/L	≥1.1 thousand cells/uL
		1	<1.1 billion cells/L and ≥0.8 billion cells/L	<1.1 thousand cells/uL and ≥0.8 thousand cells/uL
		2	<0.8 billion cells/L and ≥0.5 billion cells/L	<0.8 thousand cells/uL and ≥0.5 thousand cells/uL
		3	<0.5 billion cells/L and ≥0.2 billion cells/L	<0.5 thousand cells/uL and ≥0.2 thousand cells/uL
		4	<0.2 billion cells/L	<0.2 thousand cells/uL
Thrombocytopenia*	Platelet count	0 (normal)	≥150 billions/L	≥150 thousands/uL
		1	<150 billions/L and ≥75 billions/L	<150 thousands/uL and ≥75 thousands/uL
		2	<75 billions/L and ≥50 billions/L	<75 thousands/uL and ≥50 thousands/uL
		3	<50 billions/L and ≥25 billions/L	<50 thousands/uL and ≥25 thousands/uL
		4	<25 billions/L	<25 thousands/uL

Lipid Effects

Laboratory Test	Grade	Criteria in Système International (SI) Units	Criteria in Conventional (CN) Units
Total Cholesterol	Desirable	<5.17 mmol/L	<200 mg/dL
	Borderline high	≥5.17 mmol/L and 6.21 mmol/L	≥200 mg/dL and <240 mg/dL
	High	≥6.21 mmol/L	≥240 mg/dL
Triglycerides	Normal	<1.69 mmol/L	<150 ml/dL
	Borderline high	≥1.69 mmol/L and <2.26 mmol/L	≥150 mg/dL and <200 mg/dL
	High	≥2.26 mmol/L and <5.65 mmol/L	≥200 mg/dL and <500 mg/dL
	Very high	≥5.65 mmol/L	≥500 mg/dL

Renal Function

Laboratory Test	Grade	Criteria in Système International (SI) or Conventional Units
Elevated creatinine	0 (normal)	≤ ULN
	1	> UNL and ≤1.5x ULN
	2	>1.5x ULN and ≤ 3x ULN
	3	>3x ULN and ≤ 6x ULN
	4	>6x ULN

7. References

Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305(6846):160-164. https://doi.org/10.1136/bmj.305.6846.160

Bretz F, Pinheiro JC, Branson M. Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics*. 2005;61(3):738-748. https://doi.org/ 10.1111/j.1541-0420.2005.00344.x

Cella D, Webster K. Linking outcomes management to quality-of-life measurement. *Oncology (Williston Park)*. 1997;11(11A):232-235.

Chobanian, AV, Bakris GL, Black HR, et al; National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure; National High Blood Pressure Education Program Coordinating Committee. The seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA*. 2003;289(19):2560-2572. https://doi.org/10.1001/jama.289.19.2560

[FDA]. United States Food and Drug Administration. Qualification of the MCP-Mod Procedure. Request for Qualification of MCP-Mod as an efficient statistical methodology for model-based design and analysis of Phase II dose finding studies under model uncertainty. April 22, 2023. Accessed April 7, 2013. https://www.fda.gov/media/99313/download.

Gould AL. BMA-Mod: A Bayesian model averaging strategy for determining dose-response relationships in the presence of model uncertainty. *Biom J*. 2019;61(5):1141-1159. https://doi.org/ 10.1002/bimj.201700211

Ware JE, 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

CCI [redacted]

Signature Page for VV-CLIN-131347 v2.0

Approval	PPD Statistician 01-Dec-2024 22:36:04 GMT+0000

Signature Page for VV-CLIN-131347 v2.0