Protocol C5041005

A PHASE 2/3, TWO-PART STUDY TO EVALUATE THE EFFICACY AND LONG-TERM SAFETY WITH ORAL ETRASIMOD, 2 MG, ONCE DAILY IN ADULT PARTICIPANTS WITH MODERATE-TO-SEVERE ATOPIC DERMATITIS WITH A HISTORY OF PRIOR SYSTEMIC TREATMENT FAILURE

Statistical Analysis Plan
(SAP)

Version: 2

Date: 28 Jul 2023

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
1 15 Mar 2023	Amendment 1 07 Dec 2022	First version	NA
2 28 Jul 2023	Amendment 2 17 May 2023	1) Adjusted study objective considering inputs from US FDA. 2) Aligned with updates in the most recent protocol amendment 2. 3) Corrected typographical errors.	1) The change is: • Modified study objective from superiority hypothesis testing to estimation in section 5.1 and section 6. 2) The changes are: • Updated label of Part 2 as "Part 2 OL Safety" in various sections; • Updated Table 3 in section 2.2 to remove details of intercurrent events and to consolidate descriptions of safety endpoints; • Updated definitions of estimands E1 and E2, not to consider discontinuation of study treatment due to COVID-19/regional emergency siutations as a separate reason from other reasons in section 2.2; • Removed estimand E1b as it is captured in new definition of E1 in section 2.2 and removed E1b related descriptions in various sections; • Added provision on how to handle a subject enrolled

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			with baseline IGA < 3 as protocol deviation in stratification factor ub Section 3.4; • Removed "Physical Examination" as this CRF is not included in data collection, in Section 3.5.1; • Updated details of AESI to align with protocol amendment 2 in section 3.5.2; • Clarified how participants with multiple enrollments are handled in Section 4; • Clarified the definition of Day 1 (baseline) for participants who are randomized but not treated in section 4.1; • Updated JTC methodology under E2 in section 5.2.2.1 to reflect new estimand E2 definition; • Added one subgroup variable corresonding to expanded inclusion of participants with prior systemic corticosteriod failure only in Section 6.3; • Added description of unblinded interim

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
			analyses prior to formal IA in section 7; • Updated efficacy data set descriptions for Part 1 DB according to new estimand definitions in Appendix 3. 3) Corrected typographical errors in relevant sections including Table 5.

2. INTRODUCTION

Etrasimod is a S1PRM and is being developed as an oral treatment for patients with moderate-to-severe AD. C5041005 is a two-part study with Part 1 being 16-week, randomized, DB, placebo controlled, with an OLE, to assess efficacy and safety of 2 mg etrasimod administered orally, QD in participants with refractory, moderate-to-severe AD, whose disease is not adequately controlled with other systemic drug products, including biologics, or when use of those therapies is inadvisable (ie, participants with a history of a prior systemic treatment failure). The OL Part 2 is to assess long-term safety of etrasimod 2 mg QD in participants with moderate-to-severe AD with a history of prior systemic therapy failure. This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C5041005. In this SAP, in order to differentiate the different phases of this study, the label "Part 1 DB" is used to refer to the DB phase of Part 1, "Part 1 OLE" to the OLE phase of Part 1, and "Part 2 OL Safety" to the Part 2 of this study.

2.1. Modifications to the Analysis Plan Described in the Protocol

There are no modifications in this SAP (Version 2, 28 Jul 2023) to the analysis plan described in the protocol amendment 2 (17 May 2023).

The previous SAP version (Version 1, 15 Mar 2023) was written based on the most recent approved protocol (protocol amendment 1, 07 Dec 2022) and other protocol-related documents (eg, PACL) preceding this SAP (Section 1). This SAP was intended to provide more details on statistical methodology and analyses in accordance to the study objectives (Section 2.2). This SAP also further clarified or modified what was outlined in the protocol version where appropriate. These clarification or modification are provided in the table below.

Table 2. Clarification and Modification to Statistical Concepts or Methods Outlined in Protocol Amendment 1, 07 Dec 2022

Section of Protocol	Protocol Description	Clarification or Modification
Section 1: Protocol Summary, Section 3: Estimands	The intercurrent events for Estimand E1 and E2 are "the benefit of additional prohibited medications and regardless of treatment compliance."	The intercurrent events for Estimand E1 and E2 are updated to "initiation of prohibited medications or discontinuation of study treatment" (whichever is earlier) to reflect the intention to objectively account for these two post-randomization events for these two estimand strategies. These changes are detailed in Section 2.2 and Appendix 3.
Section 1: Protocol Summary, Section 3: Estimands	Under Exploratory Objectives for Part 1 OLE and Part 2, "Baseline, for efficacy endpoints, is defined as pre-dose Day 1 in the double-blind phase."	The definitions of baseline for Part 1 and Part 2 are different. The baseline described in protocol is for Part 1 only, since Part 2 is a separate part of the study with de novo participant enrollment. Baseline for Part 2 is the pre-dose Day 1 in the Part 2 OL phase. These changes are detailed in Section 2.2 and Section 3.4.

2.2. Study Objectives, Endpoints, and Estimand

Table 3. Objectives, Endpoints, and Estimands

Type	Objective	Endpoint	Estimand
Primary			
Efficacy	• Part 1 DB: To evaluate the efficacy of etrasimod 2 mg QD versus placebo in adult participants with moderate-to-severe AD and a history of a prior systemic therapy failure.	• Proportion of participants achieving IGA of clear (0) or almost clear (1) (on a 5-point scale) and a reduction of ≥2 points from baseline at Week 16.	• Estimand E1 (Primary Estimand): The difference in the proportions of the binary endpoint between etrasimod 2 mg QD and placebo in patients with moderate-to-severe AD and a history of a prior systemic therapy failure. More details of Estimand E1 (Primary Estimand) are described in Section 2.2.1.
Safety (depending on Part 1 DB IA results)	• Part 1 OLE and Part 2 OL Safety: To evaluate the long-term safety of oral etrasimod 2 mg QD in adult participants with moderate-to-severe AD and a history of a prior systemic therapy failure.	• Incidence and severity of treatment-emergent AEs, AEs leading to study treatment discontinuation, SAEs, and AESIs. • Incidence of clinically significant changes in clinical laboratory values,	• Not applicable.

Table 3. Objectives, Endpoints, and Estimands

Type	Objective	Endpoint	Estimand
		ECG measurements and vital signs. Note: Baseline for Part 1 is the pre-dose Day 1 in Part 1. Baseline for Part 2 is the pre-dose Day 1 in the Part 2 OL Safety phase.	
Secondary			
Efficacy	• Part 1 DB: To evaluate the efficacy of etrasimod 2 mg QD versus placebo in adult participants with moderate-to-severe AD and a history of a prior systemic therapy failure.	• Proportion of participants achieving EASI-75 at Week 16.	• Estimand E1: described above. More details of Estimand E1 are described in Section 2.2.2.
		• Percent change from baseline in EASI score at Week 16.	• Estimand E2: The mean difference in percent change from baseline in EASI score between etrasimod 2 mg QD and placebo in patients with moderate-to-severe AD and a history of a prior systemic therapy failure. More details are described in Section 2.2.2.
Exploratory			
Efficacy	• Part 1 DB: To evaluate the efficacy of etrasimod 2 mg QD versus placebo based on additional measures in adult participants with moderate-to-severe AD and a history of a prior systemic therapy failure.	• Proportion of participants achieving IGA of clear (0) or almost clear (1) (on a 5-point scale) and a reduction of ≥2 points from baseline at all timepoints except Week 16. • Proportion of participants achieving EASI-75 at all timepoints except Week 16. • Percent change from baseline in EASI score at all timepoints except Week 16.	• Estimand E1 described above for all binary endpoints. • Estimand E2 described above for all continuous endpoints.

Table 3. Objectives, Endpoints, and Estimands

Type	Objective	Endpoint	Estimand
		• Proportion of participants achieving ≥4-point reduction in PP-NRS score from baseline at all timepoints. • Proportion of participants achieving ≥4-point reduction in Skin Pain NRS score from baseline at all timepoints. • Proportion of participants achieving EASI-90 at all timepoints. • Proportion of participants achieving ≥4-point reduction in POEM from baseline at Week 16.	
Efficacy (depending on Part 1 DB IA results)	• Part 1 OLE and Part 2: To evaluate the long-term efficacy of oral etrasimod 2 mg QD in adult participants with moderate-to-severe AD and a history of a prior systemic therapy failure.	• Proportion of participants achieving EASI-75 at all scheduled timepoints. • Proportion of participants achieving IGA of clear (0) or almost clear (1) (on a 5-point scale) at all scheduled timepoints. • Percent change from baseline in EASI score at all scheduled timepoints. • Proportion of participants achieving ≥4-point reduction in PP-NRS score from baseline at all scheduled timepoints. • Proportion of participants achieving ≥4-point reduction in Skin Pain NRS score from baseline at all scheduled timepoints. • Proportion of participants achieving EASI-90 at all scheduled timepoints.	• Not applicable.

Table 3. Objectives, Endpoints, and Estimands

Type	Objective	Endpoint	Estimand
		• Change from baseline in POEM at all scheduled timepoints. Note: Baseline for Part 1 is pre-dose Day 1 in Part 1. Baseline for Part 2 is the pre-dose Day 1 in the Part 2 OL Safety phase.	

2.2.1. Primary Estimand

There are two intercurrent events considered for Part 1 DB phase of this study, (1) the initiation of prohibited medications, and (2) the discontinuation of the study treatment (etrasimod or placebo) due to any reason. Estimand E1 is an estimand strategy for a binary endpoint at a visit, when applied to the primary endpoint of IGA response at Week 16, it is known as the primary estimand. Estimand E1 is illustrated using the primary endpoint below.

Estimand E1 for a binary endpoint at a visit:

Estimand E1 uses a composite strategy and estimates the treatment effect at a visit considering a participant as non-responder after experiencing the intercurrent event of initiation of prohibited medications or discontinuation of study treatment. The primary estimand is Estimand E1 applied to the primary endpoint of IGA response at Week 16, and is defined according to the primary objective. The primary estimand is defined by the following five attributes:

1. Treatment condition: received study intervention of etrasimod 2 mg QD or placebo.

2. Population: Participants with moderate-to-severe AD as defined by the inclusion and exclusion criteria.

3. Variable: IGA response at Week 16 is score of clear (0) or almost clear (1) and a reduction from baseline of ≥2 points. A participant who experiences the intercurrent event of initiation of prohibited medications or discontinuation of study treatment due to any reason (whichever is earlier), will be considered an IGA non-responder (ie, treatment failure) at Week 16.

4. Intercurrent events:

   - Initiation of prohibited medications – a participant who receives prohibited medications post-randomization prior to the Week 16 visit will be considered an IGA non-responder (ie, treatment failure) at Week 16.

   - Discontinuation of study intervention due to any reason prior to the Week 16 visit – a participant who discontinues study treatment due to any reason will be considered an IGA non-responder (ie, treatment failure) at Week 16.

5. Population-level summary: The difference in IGA response rates at Week 16 between etrasimod 2 mg QD and placebo.

2.2.2. Secondary Estimands

Estimand E1 will be similarly applied to the binary secondary endpoint (ie, EASI-75 at Week 16) as well as the binary exploratory endpoints (eg, ≥4-point reduction from baseline in PP-NRS score at Week 16) with the endpoint substituted appropriately.

Estimand E2 strategy is defined for a continuous secondary endpoint at a visit (eg, percent change from baseline in EASI score at Week 16).

Estimand E2 for a continuous endpoint at a visit:

Estimand E2 uses a composite strategy and estimates the treatment effect at a visit considering a participant as having no treatment benefit after experiencing the intercurrent event of initiation of prohibited medications or discontinuation of study treatment. Estimand E2 is applied to the secondary endpoint of percent change from baseline in EASI score at Week 16, and is defined according to the secondary objective. This Estimand E2 is defined by the following five attributes:

1. Treatment condition: received study intervention of etrasimod 2 mg QD or placebo.

2. Population: Participants with moderate-to-severe AD as defined by the inclusion and exclusion criteria.

3. Variable: Percent change from baseline in EASI score at Week 16. If a participant in the etrasimod 2 mg QD group experiences the intercurrent events of initiation of prohibited medications or discontinuation of study treatment due to any reason (whichever is earlier), the endpoint data after the intercurrent event (Appendix 3, Post-intercurrent-event data) will be treated as missing and then imputed via multiple imputations as if they are in placebo group (ie, no treatment benefit).

4. Intercurrent events:

   - Initiation of prohibited medications – a participant who receives prohibited medications post-randomization prior to Week 16, the endpoint data after the intercurrent event (Appendix 3, Post-intercurrent-event data) will be treated as missing and then imputed via multiple imputation as if they are in placebo group (ie, no treatment benefit).

   - Discontinuation of study intervention due to any reason prior to the Week 16 visit – a participant who discontinues study treatment due to any reason, the endpoint data after the intercurrent event (Appendix 3, Post-intercurrent-event data) will be treated as missing and then imputed via multiple imputation as if they are in placebo group (ie, no treatment benefit).

5. Population-level summary: The mean difference in percent change from baseline in EASI score at Week 16 between etrasimod 2 mg QD and placebo.

Estimand E2 will also be applied to percent change from baseline in EASI score at visits prior to Week 16.

2.2.3. Additional Estimand

One supportive estimand is defined to support the Estimand E1 (primary estimand) of the primary endpoint (IGA response at Week 16) as well as Estimand E1 of IGA response at all visits prior to Week 16. It is Estimand E1a, detailed below.

Estimand E1a for IGA response at Week 16:

Estimand E1a uses a treatment policy strategy and estimates the treatment effect at a visit regardless of initiation of prohibited medications or discontinuation of study treatment. This supportive estimand is defined by the following five attributes:

1. Treatment condition: received study intervention of etrasimod 2 mg QD or placebo.

2. Population: Participants with moderate-to-severe AD as defined by the inclusion and exclusion criteria.

3. Variable: IGA response at Week 16 is score of clear (0) or almost clear (1) and a reduction from baseline of ≥2 points. The intercurrent events of initiation of prohibited medications or discontinuation of study treatment due to any reason will not be considered here. All IGA data collected including those collected after the intercurrent events are also included to derive the IGA response at Week 16.

4. Intercurrent events:
    - Initiation of prohibited medications is not considered for data exclusion.
    - Discontinuation of study intervention due to any reason is not considered for data exclusion.
5. Population-level summary: The difference in IGA response rates at Week 16 between etrasimod 2 mg QD and placebo.

The difference between Estimand E1 and E1a is that Estimand E1a disregards intercurrent events and includes all data collected regardless of intercurrent events.

One supportive estimand is defined to support the Estimand E2 of the continuous secondary endpoint at a visit (eg, percent change from baseline in EASI score at Week 16). It is Estimand E2a, detailed below.

Estimand E2a for percent change from baseline in EASI score at Week 16:

Estimand E2a uses a treatment policy strategy and estimates the treatment effect at a visit regardless of initiation of prohibited medications or discontinuation of study treatment. Estimand E2a is applied to the secondary endpoint of percent change from baseline in EASI score at Week 16. This Estimand E2a is defined by the following five attributes:

1. Treatment condition: received study intervention of etrasimod 2 mg QD or placebo.
2. Population: Participants with moderate-to-severe AD as defined by the inclusion and exclusion criteria.
3. Variable: Percent change from baseline in EASI score at Week 16. The intercurrent events of initiation of prohibited medications or discontinuation of study treatment will not be considered here. All data collected including those collected after the intercurrent events are also included in statistical analysis at Week 16.
4. Intercurrent events:
    - Initiation of prohibited medications is not considered for data exclusion.
    - Discontinuation of study intervention due to any reasons is not considered for data exclusion.
5. Population-level summary: The mean difference in percent change from baseline in EASI score at Week 16 between etrasimod 2 mg QD and placebo.

Estimand E2a will also be applied to percent change from baseline in EASI score at visits prior to Week 16.

It can be noted that Estimand E1a for binary efficacy endpoint and Estimand E2a for continuous efficacy endpoint use the same treatment-policy strategy in estimand definition. No estimands will be defined for efficacy endpoints in Part 1 OLE and Part 2 OL Safety phases, and all safety endpoints in this study, so all data collected in these phases will be used in analyses.

2.3. Study Design

This is a 2-part study with Part 1 being 16-week, randomized, DB, placebo controlled, with an OLE, to assess efficacy and safety of 2 mg etrasimod administered orally, QD in participants with refractory, moderate-to-severe AD, whose disease is not adequately controlled with other systemic drug products, including biologics, or when use of those therapies is inadvisable (ie, participants with a history of a prior systemic treatment failure). The OL Part 2 is to assess long-term safety of etrasimod 2 mg QD in participants with moderate-to-severe AD with a history of prior systemic therapy failure.

Part 1: Approximately 60 participants with moderate-to-severe AD with a history of prior systemic treatment failure will be randomized (1:1 ratio) in a double-blind manner to receive etrasimod 2 mg QD or placebo orally, once daily, for 16 weeks. Randomization will be stratified by disease severity as measured by IGA score (3 [moderate AD], 4 [severe AD]) at baseline. Following completion of the double-blind, placebo-controlled period, all participants with no clinically significant treatment-related safety concerns will be given the option to continue in an OLE whereby they will receive etrasimod 2 mg (tablet) QD for 52 weeks. Non-responders, participants not achieving EASI-50, will be discontinued at Week 32 (ie, 16 weeks of treatment in the OLE period).

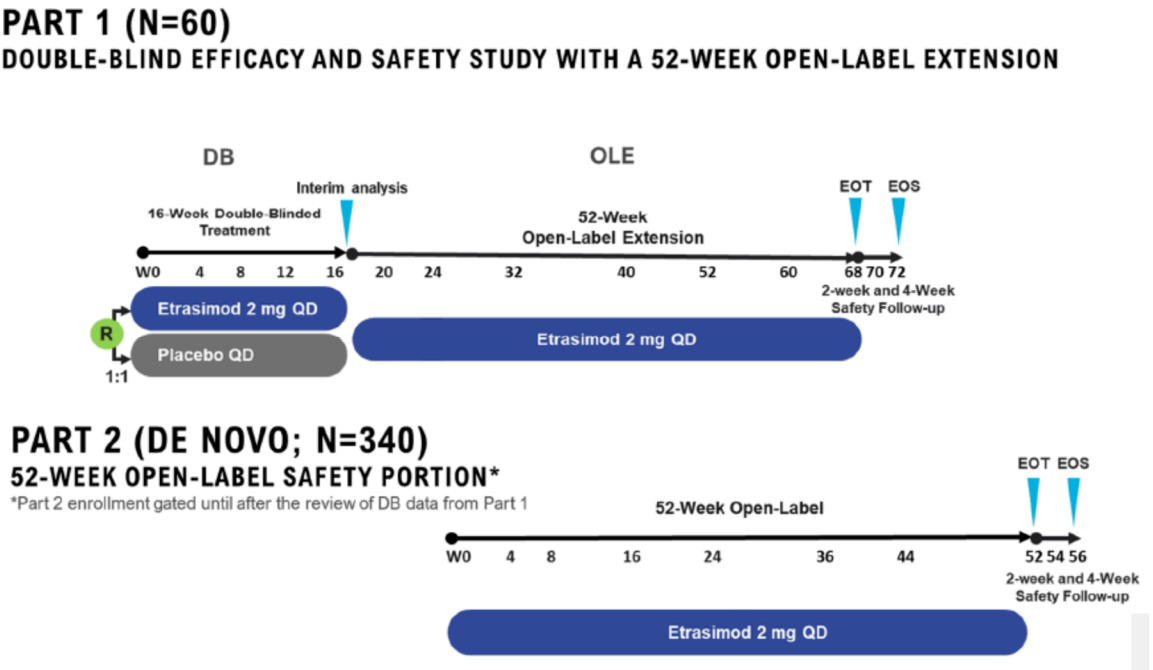
Part 2: An interim analysis will be conducted after all applicable Part 1 participants have completed the 16-week DB treatment period. The IA will be performed by the Sponsor and data summaries will be shared with the study team. The study team will review the data in the context of the overall benefit-risk profile of etrasimod for the treatment of moderate-to-severe AD with prior systemic failure. The data review will include but is not limited to, efficacy measures of IGA and EASI-75 response and safety parameters including the incidence and severity of AEs, AESIs, and clinically relevant changes in ECG, and vital signs, and laboratory values. The efficacy and safety data will be utilized to make a go/no-go decision for initiating Part 2 of the study, as well as for completing the 52-week treatment for the Part 1 participants in the OLE. Once the decision is made it will be promptly communicated to the sites.

If the IA results in a decision to proceed to Part 2 of the study, approximately 340 additional (de novo) participants will be enrolled to receive etrasimod 2 mg orally, once daily, for 52 weeks in an open-label manner in order to fulfill the safety database requirements. Enrollment into Part 2 will not begin until the IA data from Part 1 DB are reviewed, and determination made as to the favorable benefit/risk of etrasimod 2 mg QD in this new target

population. Non-responders, ie, participants not achieving at least EASI-50 will be discontinued at Week 16, Part 2. The participant population in both Part 1 and Part 2 will include at least 40% of participants with moderate AD (IGA score of 3) and at least 40% of participants with severe AD (IGA score of 4). Figure 1 below provides a schematic of this study design.

The application of topical emollients/moisturizers (including those containing ceramide, hyaluronic acid, or urea) will be required at least once daily for at least 1 week prior to baseline (Day 1) and throughout the treatment period without change (ie, type, frequency, application). TCS and other medicated topical treatments for AD must be discontinued at least 1 week prior to baseline (Day 1) and are not permitted in the double-blind treatment period. Safety follow-up visits will occur at 2 and 4 weeks after the last dose of study treatment. Medicated and non-medicated rescue medications are not permitted during the Part 1 DB portion. For Part 1 OLE and Part 2 medicated and non-medicated topical treatments for AD will be permitted at the discretion of the investigator and in accordance with their usual practice.

Figure 1. Study Schematic



3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoints

For Part 1 DB,

- IGA response at Week 16.

  IGA response is defined as clear (0) or almost clear (1) on a 5-point scale and a reduction of ≥2 points from baseline (Appendix 2.2.1). Only participants with baseline IGA ≥2 in the FAS are included in all analyses of IGA response.

For Part 1 OLE and Part 2,

- Incidence and severity of treatment-emergent AEs, AEs leading to study treatment discontinuation, SAEs, and AESIs.
- Incidence of clinically significant changes in clinical laboratory values, ECG measurements and vital signs.

3.2. Secondary Endpoints

For Part 1 DB,

- EASI-75 at Week 16.

  EASI-75 response is defined as ≥75% reduction in EASI from baseline (Appendix 2.2.3).

- Percent change from baseline in EASI score at Week 16 (Appendix 2.2.3).

3.3. Exploratory Endpoints

For Part 1 DB,

- IGA response at at all timepoints except Week 16.
- EASI-75 at all timepoints except Week 16.
- Percent change from baseline in EASI score at all timepoints except Week 16.
- ≥4-point reduction in PP-NRS score from baseline at all timepoints (Appendix 2.2.4). Only participants with baseline PP-NRS ≥4 in the FAS are included in all analyses of this endpoint.
- ≥4-point reduction in Skin Pain NRS score from baseline at all timepoints (Appendix 2.2.5). Only participants with baseline Skin Pain NRS ≥4 in the FAS are included in all analyses of this endpoint.

- EASI-90 at all timepoints.

  EASI-90 response is defined as ≥90% reduction in EASI from baseline (Appendix 2.2.3).

- ≥4-point reduction in POEM from baseline at Week 16 (Appendix 2.2.6). Only participants with baseline POEM ≥4 in the FAS are included in all analyses of this endpoint.

For Part 1 OLE and Part 2,

- EASI-75 at all scheduled timepoints.
- IGA of clear (0) or almost clear (1) on a 5-point scale at all scheduled timepoints.

  It is important to note that this is different from IGA response, in that this is without the ≥2-point reduction component.

- Percent change from baseline in EASI score at all scheduled timepoints.
- ≥4-point reduction in PP-NRS score from baseline at all scheduled timepoints. Only participants with baseline PP-NRS ≥4 are included in all analyses of this endpoint.
- ≥4-point reduction in Skin Pain NRS score from baseline at all scheduled timepoints. Only participants with baseline Skin Pain NRS ≥4 are included in all analyses of this endpoint.
- EASI-90 at all scheduled timepoints.
- Change from baseline in POEM at all scheduled timepoints.

3.4. Baseline Variables

For Part 1 DB and OLE, baseline value for efficacy and safety endpoints is defined as pre-dose value obtained on Day 1 (ie, baseline visit) or prior (if the Day 1 value is not available) in the Part 1 double-blind phase in the analyses. For Part 2, baseline value is defined as pre-dose value obtained on Day 1 (ie, baseline visit) or prior (if the Day 1 value is not available) in the Part 2 open-label phase in the analyses.

A participant's baseline IGA score value (3 [moderate AD], 4 [severe AD]) recorded in the clinical database will be used as the actual IGA stratification factor for both Part 1 and Part 2. If a participant is randomized but the actual baseline IGA is less than 3 (ie, protocol deviation), this participant will be grouped in the actual stratum of baseline IGA score of 3, in order to keep the participant in analyses. This is described in more detail below.

3.4.1. Stratification Factor

Participants will be stratified by disease severity as measured by baseline IGA score (3 [moderate AD], 4 [severe AD]). The stratum derived from data collected on IGA score at baseline will be used for all analyses requiring this stratification variable (Section 5.2). This stratification is called the actual IGA stratification variable. The stratum of IGA score reported to the randomization system is called the reported IGA stratification.

For the primary endpoint of IGA response at Week 16 in Part 1 DB phase, a sensitivity analysis will be performed using the reported IGA stratification if there are discrepancies between the actual IGA stratification and the reported IGA stratification. A participant listing will be provided comparing the reported IGA stratification variable (to the randomization system) and actual IGA stratification variable (derived using the clinical database).

3.5. Safety Endpoints

All the safety data will be summarized descriptively through appropriate data tabulations, descriptive statistics, and graphical presentations. Treatment-emergent AEs including SAEs and AE leading to discontinuation of study treatment, vital signs, 12-lead ECG parameters, and laboratory tests, etc. will be summarized according to the Clinical Data Interchange Standards Consortium (CDISC) and Pfizer Standards (CaPS). Treatment-emergent AE is defined as an AE that started on or after the first dose of study treatment (ie, baseline, without regard to baseline severity). The MedDRA version for AE coding at the time of database release will be adopted and the version used will be displayed in all relevant safety data tabulations. All safety analyses will be performed on the SAS (Section 4.2).

3.5.1. Adverse Events (Including Laboratory Data)

In this study, the safety endpoints included in safety data summaries are listed in, although not limited to, the following:

- Incidence and severity of treatment-emergent AEs, AE leading to study treatment discontinuation, SAEs, and AESI, etc.
- Incidence of clinically significant changes in clinical laboratory values, 12-lead ECG measurements, and vital signs, etc.
- Change from baseline in laboratory values (hematology, serum chemistry, coagulation, and urinalysis), 12-lead ECG measurement, vital signs, and other objective safety measures.

These safety endpoints are considered as primary endpoints for the Part 1 OLE and Part 2 OL Safety phases. Definitions of baseline for Part 1 and Part 2 are previously provided (Section 3.4).

3.5.2. Adverse Events of Special Interest

Treatment-emergent AESIs are identified per study protocol and require close monitoring. All AESIs must be reported as an AE or SAE following the reporting procedures described in

the protocol. An AESI that is also an SAE must also be reported using the CT SAE Report Form. The following are AESI identified for etrasimod (except for PML, the MedDRA search terms for AESI are recorded in a separate charter for the etrasimod clinical program):

- Cardiovascular events (eg, bradycardia, AV conduction delay, and hypertension);
- Macular oedema*;
- Pulmonary events (airflow obstruction or altered gas exchange);
- Infections (severe infections, opportunistic infections including PML*, herpes simplex and herpes zoster);
- Liver injury (liver transaminase elevation and bilirubin elevation);
- PRES*; and
- Malignancies*.

* A diagnosis of this event requires permanent discontinuation of study intervention.

3.5.3. Study Treatment Compliance

In addition to safety endpoints, the study treatment compliance (%) will be derived from the total number of doses actually taken divided by the total number of doses expected to take between two scheduled visits using dosing data collected in the Treatment Dosing CRF page. Compliance is expressed in percentage. Both the randomized treatments (ie, etrasimod 2 mg QD and placebo) will be taken into account. Participants who have study treatment compliance of <80% or >120% between two scheduled visits will be identified in a listing.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

Data for all participants will be assessed to determine if participants meet the criteria for inclusion in each analysis set (population) prior to unblinding, releasing the database for analyses. Classification of participants to analysis sets will be documented per standard operating procedures. For participants with multiple enrollments, only the participant identifier and data under initial enrollment will be included and summarized. Data collected for other enrollments with be included only in listings and narratives.

Participants who are screened but do not meet inclusion criteria or meet exclusion criteria to enter the study (ie, screen failure), they will be counted as a single group and this is how far they are summarized (Section 6.4.2). Below are descriptions of the analysis sets defined for this study.

4.1. Full Analysis Sets

The full analysis set (FAS) will include all participants who are randomized to the study irrespective of whether they receive any dose of study intervention (ie, etrasimod or placebo). Participants will be analyzed in the treatment groups as they are randomized. If a participant

is treated but not randomized, then this participant will be excluded from the FAS. A narrative will be provided for this participant in the clinical study report (CSR). For a participant who is randomized but not treated, for Part 1 DB and OLE, baseline value for efficacy and safety endpoints is defined as the value obtained on the date of randomization as Day 1 or prior (if the Day 1 value is not available) in the Part 1 DB phase in the analyses.

In general, the analysis population for analzying all efficacy endpoints in this study is defined by the complete FAS. However, there are some exceptions: some efficacy endpoints, defined by specific point reduction from baseline, require the inclusion of only participants who meet a baseline threshold in the analysis set to allow them the potential to respond in these efficacy endpoints. The details of these exceptions are provided below. Therefore, when considering analyzing an efficacy endpoint under an estimand, it is important to consider these exceptions.

FAS for Specific Efficacy Endpoints

The following four binary efficacy endpoints will only be analyzed including participants who meet the specific baseline threshold. This is to ensure that only participants with the potential to respond at a post-baseline visit are included in the analyses.

- IGA response defined as clear (0) or almost clear (1) on a 5-point scale and a reduction of ≥2 points from baseline (Appendix 2.2.1): Only participants with baseline IGA ≥2 in the FAS are included in all analyses of IGA response, knowing that only participants with baseline IGA ≥3 are eligible to enroll in this study.

- ≥4-point reduction in PP-NRS score from baseline (Appendix 2.2.4): Only participants with baseline PP-NRS ≥4 in the FAS are included in all analyses of this endpoint.

- ≥4-point reduction in Skin Pain NRS score from baseline (Appendix 2.2.5): Only participants with baseline Skin Pain NRS ≥4 in the FAS are included in all analyses of this endpoint.

- ≥4-point reduction in POEM from baseline (Appendix 2.2.6): Only participants with baseline POEM ≥4 in the FAS are included in all analyses of this endpoint.

For all other efficacy endpoints, the general rule of including the complete FAS still applies.

4.2. Safety Analysis Set

The safety analysis set (SAS) will include all participants who are randomized and receive at least one dose of the study treatment (ie, etrasimod or placebo). Participants will be analyzed in the treatment groups as they receive. If a participant is treated but not randomized, then the participant will be excluded from SAS. A narrative will be provided for this participant in the CSR. The SAS will be the analysis set for safety analyses in all phases of the study.

5. GENERAL METHODOLOGY AND CONVENTIONS

For Part 1, there will be a total of 2 planned analyses. The first (interim) analysis (IA) for Part 1 will be conducted when all applicable participants have completed or discontinued from study prior to their Part 1 DB Week 16 visit. The IA will be performed by the Sponsor and it will include all efficacy and safety data through Week 16. Data summaries from the IA will be shared with the study team. The study team will review the data in the context of the overall benefit-risk profile of etrasimod for the treatment of moderate-to-severe AD with prior systemic failure. The efficacy and safety data up to Week 16 will be utilized to assess the primary efficacy endpoint, secondary efficacy endpoints, and safety endpoints (Section 2.2, Section 5.1) and to make a go/no-go decision for initiating Part 2 of the study, as well as for completing the 52-week treatment for the Part 1 participants in the OLE. The IA efficacy results through Week 16 for Part 1 DB phase is not final and definitive.

The second (final) analysis for Part 1 will be conducted when all applicable participants have completed or discontinued from study prior to their Week 68 visit (including safety follow-up visits at Weeks 70 and 72) if Part 1 OLE is planned for completion by the sponsor. Otherwise, if Part 1 OLE is decided to be terminated by the sponsor, the final analysis will only be conducted based on data collected up to the date of study termination. The final analysis will contain final safety and efficacy data, as Part 1 final database, to assess the primary endpoints, secondary endpoints as well as exploratory efficacy endpoints (Section 2.2). This will also include efficacy analyses for Part 1 DB phase repeated based on this final data, and this final analysis is considered final and definitive.

For Part 2, there will only be one final analysis when all applicable participants have completed their Part 2 OL Safety phase Week 52 visit (including safety follow-up visits at Weeks 54 and 56) if Part 2 OL Safety phase is planned by the sponsor based on the Part 1 DB IA result. The final analysis will contain final safety data to assess the primary safety endpoints, and final efficacy data to assess exploratory efficacy endpoints (Section 2.2).

5.1. Hypotheses and Decision Rules

This protocol is designed to estimate the treatment difference in response rates of etrasimod 2 mg QD versus placebo, as well as their individual response rates (Section 5.2.1.1), for the treatment of moderate-to-severe AD based on the primary endpoint of IGA response rate at Week 16 in Part 1 DB (based on Estimand E1 as primary estimand) among adult participants (with baseline IGA ≥2) with a history of a prior systemic therapy failure. There are no statistical hypotheses in this study, therefore no multiplicity adjustment will be made in this study.

5.2. General Methods

In Part 1 DB, as all participants randomized to treatment groups of etrasimod 2 mg QD and placebo will switch to open-label treatment of etrasimod 2 mg QD at Week 16, the treatment labels used for reporting DB visits up to Week 16 will be “Etrasimod 2 mg QD” and “Placebo” respectively. At the end of Part 1 OLE, when reporting OLE visits from Week 20 through Week 68, the treatment labels used will be “Etrasimod 2 mg QD” and “Placebo → Etrasimod 2 mg QD” respectively to differentiate these two treatment groups.

For Part 1 DB analyses up to Week 16, the following treatment comparison will be made at each timepoint, where applicable:

- Etrasimod 2 mg QD vs. Placebo.

The primary efficacy comparison for the primary endpoint of IGA response at Week 16 will be between etrasimod 2 mg QD and placebo. For analyses after Week 16 from Week 20 through Week 68 (Part 1 OLE phase), the following treatment comparison will be made at each timepoint,

- Etrasimod 2 mg QD vs. Placebo → Etrasimod 2 mg QD.

For Part 2 OL Safety phase, there will only be one treatment group of etrasimod 2 mg QD, the label will be "Etrasimod 2 mg QD". There is no treatment comparison in Part 2. There will be no pooling of efficacy and safety data of Part 1 and Part 2 in this study.

Descriptive Summaries

In general, the data for all continuous endpoints will be summarized by treatment group and by timepoint in tables containing descriptive statistics (N as the number of participants evaluable for the endpoint at the timepoint, mean, standard deviation, standard error of the mean, minimum, 1st, 2nd (ie, median) and 3rd quartiles and maximum) for actual and change from baseline (or percent change from baseline) values for those endpoints measured at baseline. In case when N=1 (ie, only one particpant is available/evaluable for summary), standard deviation and standard error will be reported as "NA" (not applicable) while the remaining statistical parameters will have the same value as the mean. The data for all response-type binary endpoints will be summarized by treatment group and by timepoint in tables showing descriptive statistics: N, n (ie, number of responders), response rate, standard error (SE) of the response rate, and 95% confidence interval (CI) based on normal approximation, all expressed in %. If $\hat{p}$ is the estimated response rate, then the 95% CI is calculated as:

$$\hat{p} \pm z_{0.975} SE = \hat{p} \pm z_{0.975} \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$$

where $z_{0.975}$ is the 97.5th percentile of the standard normal distribution and N is the number of participants evaluable for the endpoint at the timepoint. If the lower bound is calculated to be negative, it will be set to 0%; if the upper bound is calculated to be larger than 100%, it will be set to 100%. In case when response rate is 0 or 100%, standard error will be reported as "NA" and the 95% CI bounds will be the same as the response rate. The displays described above for continuous and response-type binary endpoints will only use available data with no imputation. Therefore, the calculation of response rates will use the number of evaluable participants as denominators.

5.2.1. Analyses for Binary Endpoints

5.2.1.1. Cochran-Mantel Haenszel (CMH) Approach

At a single timepoint (eg, IGA response at Week 16 in Part 1 DB), difference in response rate between etrasimod 2 mg QD and placebo groups will be estimated using the Cochran-Mantel Haenszel (CMH) approach adjusting for the stratification factor of IGA score (3 [moderate AD], 4 [severe AD]) at baseline) (Cochran, 1954; Mantel and Haenszel, 1959). Large sample approximation will be used for testing the superiority of etrasimod 2 mg to placebo at Part 1 Week 16 and for forming 95% CI and calculating p-value. This approach will be referred to as CMH. This method is also applicable for all other timepoints in Part 1.

Explicitly, let $\hat{p}_{Ai}$ and $\hat{p}_{Bi}$ represent the estimated response rates for etrasimod 2 mg (A) and placebo (B), respectively, of the i^{th} stratum. Using the CMH weights, the weighted difference in rates between etrasimod 2 mg QD and placebo groups, d, is expressed by:

$$\text{d} = \sum_i w_i (\hat{p}_{Ai} - \hat{p}_{Bi})$$

where $w_i = \frac{(n_{Ai} n_{Bi})/(n_{Ai}+n_{Bi})}{\sum_i (n_{Ai} n_{Bi})/(n_{Ai}+n_{Bi})}$, is the CMH weight for the i^{th} stratum, with n_{Ai} and n_{Bi} equal the numbers of participants in the etrasimod and placebo groups, respectively, per stratum. The weights are normalized such that the sum across strata adds up to 1.

Two-sided 95% CI will be estimated using the normal approximation to the binomial distribution via the following method,

$$d \pm \left[z_{1-\alpha/2} \sqrt{var(d)} \right]$$

where the variance of d is var(d) = $\sum_i w_i^2$ var($\hat{p}_{Ai} - \hat{p}_{Bi}$) with

$$\text{var}(\hat{p}_{Ai} - \hat{p}_{Bi}) = \frac{\hat{p}_{Ai}(1-\hat{p}_{Ai})}{n_{Ai}} + \frac{\hat{p}_{Bi}(1-\hat{p}_{Bi})}{n_{Bi}}.$$

Two-sided p-value for the test of the 0 difference between etrasimod 2 mg QD and placebo groups will be calculated as:

$$p = 2(1 - \Phi(|Z|)),$$

where $\Phi(.)$ is the Gaussian cumulative density function and $Z = \frac{d}{\sqrt{\text{var}(d)}}$ is the Normal Z test statistic.

If there is no (ie, 0%) response or 100% response rate in any one or both of the two treatment groups for the comparison in a stratum, eg, etrasimod vs. placebo, when calculating the proportions above, 0.5 will be added to the number of responses (ie, numerator) and 1 will be

added to the denominator in each treatment in that stratum only corresponding to the pair of comparison for calculating the treatment difference, standard error (ie, $\sqrt{var(d)}$), 95% CI and 2-sided p-value (Agresti, 2002).

When response rate of 0% or 100% is observed in both treatments in comparison and in both strata, no formal comparison will be performed. Estimated response rate of 0% or 100% will be reported as observed. Standard error will be reported as 0. The final results will be expressed in percentages, ie, (proportions x 100)%.

5.2.2. Analyses for Continuous Endpoints

5.2.2.1. Analyses for Continuous Data (Repeated Measures Under Estimand E2)

For a continuous efficacy endpoint analyzed under the Estimand E2 and measured at multiple post-baseline visits (ie, percent change from baseline in EASI score at Week 16), this endpoint will be analyzed using the Jump-to-Control (JTC) method (Carpenter, Roger and Kenward 2013). This JTC method is implemented in the following steps.

1. A mixed model for repeated measures (MMRM) that includes fixed effects of treatment group, visit, treatment group-by-visit interaction, actual stratification factor (ie, baseline IGA of 3 or 4 from clinical database), actual-stratification-factor-by-visit interaction, baseline value, and baseline-value-by-visit interaction, will be used to analyze the endpoint's data across all visits (ie, all visits up to Week 16 for Part 1 DB phase). This MMRM will include all relevant observed values across all visits in the pre-intercurrent-event-data without imputation of missing values on the FAS for Estimand E2 (Section 4.1, Appendix 3). The data after intercurrent events of initiation of prohibited medications or discontinuation of study treatment due to any reason will be set to missing (see Appendix 3 for more details on data inclusion). A common unstructured variance-covariance matrix will be used. Estimation of model parameters is performed under the Bayesian framework using MCMC method with non-informative prior densities for the model parameters. This is known as the estimation model.

2. Multiple imputation allows for uncertainty of the imputed values. Using the estimated model specified in step 1, missing percent change from baseline value at a particular visit of interest (eg, Week 16) for a participant can be imputed multiple times using multivariate conditional Normal distribution as the imputation model constructed from the parameter values sampled from their posterior densities (using MCMC) conditional on participant's observed values at prior visits, if any.

3. In order to apply the JTC concept, for a participant in the etrasimod 2 mg QD group with monotone missing values, the means of the multivariate conditional Normal distribution for those visits with monotone missing values will be equal to those means estimated for the placebo group for the same visits. The visit of interest will be subject to multiple imputation based on this distribution. This will result in multiple imputed datasets at this visit of interest. Intermittent missing values will be imputed under MAR. Participants in the placebo group with missing values at this

visit of interest will be imputed based on the MAR assumption. The number of imputation (R) can be specified as 100 here.

4. For each of the R completed imputed dataset for a visit of interest, an analysis of covariance (ANCOVA) model which includes fixed effects of treatment group, actual stratification factor (ie, baseline IGA of 3 or 4 from clinical database), and baseline value, will be used to analyze the data (observed and imputed) at the visit of interest. This will generate LSM (SE) for each treatment group and LSM treatment difference estimate as well as standard error of this estimate. This will produce R sets of estimates and standard errors. These R sets of results will be combined using the Rubin's rules (Rubin 1987) to generate the combined LSM (SE) for each treatment group and LSM treatment difference estimate and its standard error, 95% CI and 2-sided p-value, for this visit of interest.

The implementation of this JTC method for continuous endpoint measured at multiple post-baseline visits under the Estimand E2 will provide analysis result for each visit of interest. The implementation may be using the Statistical Analysis System (SAS) macro developed by Carpenter, Roger and Kenward, 2013. Further programming details of JTC method will be laid out in the programming plan.

5.2.2.2. Analyses for Continuous Data (Repeated Measures With No Estimand)

When Estimand E2a or no estimand is defined for a continuous efficacy endpoint, repeated measures data for continuous endpoints will be analyzed as change or percent change from baseline as appropriate with a MMRM that includes fixed effects of treatment group, visit, treatment group-by-visit interaction, actual stratification factor (ie, baseline IGA of 3 or 4 from clinical database), actual-stratification-factor-by-visit interaction, baseline value, and baseline-value-by-visit interaction. This MMRM is similar to the estimation model used in JTC method (Section 5.2.2.1). MMRM will include observed values without imputation of missing values. A common unstructured variance-covariance matrix will be used, provided the model converges, otherwise alternative covariance structures will be attempted in this order depending on model convergence and feasibility: heterogeneous compound symmetry (CSH) then compound symmetry (CS). The Kenward-Roger degrees of freedom approximation will be used. Comparison of treatment groups (providing LSM of the treatments, LSM of the treatment difference, 2-sided p-value and 95% CI) at each included timepoint will be generated using this MMRM. MMRM can only include participants with observed baseline value and at least one observed post-baseline value. If the baseline value is missing or if there are no post-baseline measurements, this participant will not be included this analysis (Section 4.1).Therefore, when reporting the result of MMRM, the number of participants by treatment group in the FAS (N), the number of participants by treatment group included in the MMRM (N1) and the number of participants by treatment group evaluable at each of the timpoints (N2) are to be reported.

5.3. Methods to Manage Missing Data

In Part 1 DB phase, after accounting for relevant intercurrent events in the data sets (Section 2.2, Appendix 3) and visit windowing (Appendix 2.1) in the data, if a binary efficacy endpoint at a visit has missing value, it will be handled by setting the missing binary efficacy

endpoint to non-response. This method of handling missing binary efficacy endpoint is known as missing response as non-response (MR=NR). In general, if a continuous efficacy endpoint at a visit has missing value, it will remain missing and no imputation of this missing value will be performed in the datasets.

Under Estimand E1 (eg, IGA response at Week 16), after accounting for the occurrence of relevant intercurrent events in the dataset and in the FAS, missing binary efficacy endpoint at a visit will be set to non-response (ie, treatment failure). If intercurrent events have not occurred before a visit, and the binary efficacy endpoint at this visit has missing value, it will be set to non-response due to MR=NR. Under the Estimand E1a, since intercurrent events are not considered, if the binary efficacy endpoint at a visit has missing value, it will be set to non-response due to MR=NR.

For continuous efficacy endpoint analyzed under Estimand E2 (eg, percent change from baseline in EASI score at a visit), after accounting for the occurrence of relevant intercurrent events, the post-intercurrent-event data will be treated as missing, only pre-intercurrent event data will be included in the JTC method via multiple imputations (Section 5.2.2.1).

In Part 1 OLE phase, since no estimands are defined for (binary or continuous) efficacy endpoints, all observed data collected/assessed regardless of any prior intercurrent events will be included in deriving the efficacy endpoints at a visit. If the binary efficacy endpoint at any visit has missing value, it will be handled by setting the binary efficacy endpoint to non-response. This method of handling missing binary efficacy endpoint is known as missing response as non-response (MR=NR), which is similar to using Estimand E1a (treatment policy estimand). If the continuous efficacy endpoint at any visit has missing, it will remain as missing, without imputation of missing value at this visit, which is similar to using Estimand E2a (treatment policy estimand). The same approach of including all observed data collected/assessed regardless of any prior intercurrent events will also be applicable to Part 2 OL Safety phase. Missing binary efficacy endpoints in Part 2 OL Safety phase will not be imputed as non-response, only observed values will be summarized.

In general, missing values in any (binary or continuous) efficacy endpoint will not be imputed when summarizing these endpoints using descriptive statistics. In addition, missing values for safety endpoints will not be imputed.

6. ANALYSES AND SUMMARIES

6.1. Primary Endpoint

6.1.1. Primary Analysis

The IGA response at Week 16 is the primary efficacy endpoint in Part 1 DB phase. The analysis of the primary endpoint will be based on Estimand E1 accounting for the intercurrent events using the FAS analyzed for participants with baseline IGA ≥2 (Section 4.1, Appendix 3). This corresponds to the primary estimand including only the pre-intercurrent-event data (Appendix 3). For the primary analysis, the normal approximation for the difference in binomial proportions adjusting for the actual stratification factor of baseline IGA (3 vs 4) via the CMH approach (Section 5.2.1.1) will be used. The baseline IGA from

the clinical database will be used as the actual stratification factor (Section 3.4.1). This CMH approach will be used to estimate the treatment difference in response rates of etrasimod 2 mg QD versus placebo and to generate 95% CI for the IGA response rate difference at Week 16 along with the 2-sided p-value (Section 5.2.1.1). Missing values will be handled by setting the IGA response value to non-response, MR=NR (Section 5.3). This is the primary analysis for IGA response at Week 16.

Analysis of IGA response at other visits prior to Week 16 (ie, exploratory endpoints) will also be analyzed using the same method based on the Estimand E1 including only the pre-intercurrent-event data on the FAS analyzed for participants with baseline IGA ≥2.

6.1.2. Supportive Analysis

A supportive analysis will be conducted for the primary efficacy endpoint of IGA response at Week 16 based on Estimand E1a without regard to intercurrent events using the FAS analyzed for participants with baseline IGA ≥2 (Section 2.2.3, Appendix 3). This corresponds to the supportive estimand including all data (Appendix 3). For this supportive analysis, the same normal approximation for the difference in binomial proportions adjusting for the actual stratification factor of baseline IGA (3 vs 4) via the CMH approach (Section 5.2.1.1) will be used. The baseline IGA from the clinical database will be used as the actual stratification factor (Section 3.4.1). This CMH approach will be used to support the primary analysis of estimating the treatment difference in response rates of etrasimod 2 mg QD versus placebo and to generate 95% CI for the IGA response rate difference at Week 16 along with the 2-sided p-value. Missing values will be handled by setting the IGA response value to non-response, MR=NR (Section 5.3). This is the first supportive analysis for IGA response at Week 16.

Supportive analysis of IGA response at other visits prior to Week 16 will also be performed using the same method based on the Estimand E1a including all data on the FAS, all analyzing participants with baseline IGA ≥2.

When there are discrepancies between actual stratification factor (baseline IGA of 3 vs 4) versus reported stratification factor from the randomization system, a sensitivity analysis for the IGA response at Week 16 and visits prior to Week 16 will be conducted using the normal approximation for the difference in binomial proportions adjusting for the reported (instead of actual) stratification factor of baseline IGA of 3 vs 4 via the CMH approach (Section 5.2.1.1). This includes the pre-intercurrent-event data under Estimand E1 on the FAS analyzed for participants with baseline IGA ≥2 (Section 3.4.1, Section 4.1, Appendix 3).

In addition, for the primary endpoint of IGA response at Week 16 under Estimand E1, a summary similar to the following will also be generated to assess the impact of intercurrent events on the estimates.

Table 4. IGA Response at Week 16 and Intercurrent Events

IGA Response at Week 16 – Estimand E1, FAS with Baseline IGA ≥2 for Estimand E1, MR=NR		
n (%)	Etrasimod 2 mg QD (FAS N=xxx) [1]	Placebo (FAS N=xxx) [1]
IGA Responders	xx (xx.xx)	xx (xx.xx)
IGA Non-Responders	xx (xx.xx)	xx (xx.xx)
1) Week 16 Completers with Observed Non-Response	xx (xx.xx)	xx (xx.xx)
2) Week 16 Completers with Missing Response Set as Non-Response	xx (xx.xx)	xx (xx.xx)
3) Initiation of Prohibited Medications Prior to Week 16 Visit (Earlier)	xx (xx.xx)	xx (xx.xx)
4) Discontinued Study Treatment Due to Any Reason Prior to Week 16 Visit (Earlier)	xx (xx.xx)	xx (xx.xx)
Reasons for Discontinuation of Study Treatment [2]		
Lack of Efficacy	xx (xx.xx)	xx (xx.xx)
Adverse Event	xx (xx.xx)	xx (xx.xx)
Death	xx (xx.xx)	xx (xx.xx)
Lost to Follow-up	xx (xx.xx)	xx (xx.xx)
Withdrawal by Subject	xx (xx.xx)	xx (xx.xx)
Non-Compliance with Study Drug	xx (xx.xx)	xx (xx.xx)
Protocol Deviation	xx (xx.xx)	xx (xx.xx)
Pregnancy	xx (xx.xx)	xx (xx.xx)
Physician Decision	xx (xx.xx)	xx (xx.xx)
Medication Error without Associated Adverse Event	xx (xx.xx)	xx (xx.xx)
Study Terminated by Sponsor	xx (xx.xx)	xx (xx.xx)
Other	xx (xx.xx)	xx (xx.xx)

[1] N = number of participants in the FAS with Baseline IGA ≥2. IGA response rates are calculated based on this as denominator.
[2] Only reasons with at least one participant in any group will be summarized.

6.2. Secondary and Exploratory Endpoints

6.2.1. Secondary Binary Endpoints

The EASI-75 response at Week 16 is the binary secondary efficacy endpoint in Part 1 DB phase. The analysis of the binary secondary endpoint will be based on Estimand E1 accounting for the intercurrent events using the FAS (Section 2.2.2, Appendix 3). This corresponds to the Estimand E1 including only the pre-intercurrent-event data (Appendix 3). For the secondary analysis, the normal approximation for the difference in binomial proportions adjusting for the actual stratification factor of baseline IGA (3 vs 4) via the CMH approach (Section 5.2.1.1) will be used. The baseline IGA is from the clinical database will be used as the actual stratification factor (Section 3.4.1). This CMH approach will be used to estimate the treatment difference in response rates of etrasimod 2 mg QD versus placebo and to generate 95% CI for the EASI-75 response rate difference at Week 16 along with the 2-sided p-value (Section 5.2.1.1). Missing values will be handled by setting the EASI-75

response value to non-response, MR=NR (Section 5.3). This is the secondary analysis for EASI-75 response at Week 16.

Analysis of EASI-75 response at other visits prior to Week 16 (ie, exploratory endpoints) will also be analyzed using the same CMH method based on the Estimand E1 including pre-intercurrent-event data on the FAS.

Supportive analyses for EASI-75 for visits up to Week 16 will also be performed under the Estimand E1a using FAS (including all data). The same CMH approach adjusting to actual stratification factor (Section 3.4.1) will also be used. These analyses are to support the secondary analysis of estimating the treatment difference in response rates of etrasimod 2 mg QD versus placebo for EASI-75 at Week 16.

6.2.2. Exploratory Binary Endpoints

Analysis of other exploratory endpoints for Part 1 DB phase: ≥4-point reduction from baseline in PP-NRS at all visits up to Week 16 (analyzed only for participants with baseline PP-NRS ≥4), ≥4-point reduction from baseline in Skin Pain NRS at all visits up to Week 16 (analyzed only for participants with baseline Skin Pain NRS ≥4), EASI-90 response at all visits up to Week 16, and ≥4-point reduction from baseline in POEM at Week 16 (analyzed only for participants with baseline POEM ≥4), will also be analyzed using the same CMH method based on the Estimand E1 including pre-intercurrent-event data on the FAS. No supportive analyses are performed for these exploratory endpoints.

6.2.3. Secondary Continuous Endpoint (Repeated Visits)

The percent change from baseline in EASI at Week 16 is the continuous secondary efficacy endpoint in Part 1 DB phase. The analysis of the continuous secondary endpoint will be based on Estimand E2 accounting for the intercurrent events using the FAS (Section 2.2.2, Appendix 3). This corresponds to the Estimand E2 including only the pre-intercurrent-event data (Appendix 3). For this secondary analysis, the JTC method will be used to analyze percent change from baseline in EASI score at Week 16 (Section 5.2.2.1 for details on how to handle intercurrent events). This analysis will provide the LSM of the treatment group, the LSM of the treatment difference (Etrasimod 2 mg QD versus Placebo) in percent change from baseline in EASI score at Week 16, 95% CI and the 2-sided p-value from the Rubin's rules to estimate the treatment difference of etrasimod 2 mg QD versus placebo of this continuous secondary endpoint (Section 5.2.2.1).

Analysis results of percent change from baseline in EASI score at other visits prior to Week 16 (ie, exploratory endpoints) will also be obtained from the same JTC method performed above under the same Estimand E2 using the FAS.

A supportive analysis will be performed for percent change from baseline in EASI score at visits up to Week 16 under the Estimand E2a regardless of intercurrent events using the FAS (Section 2.2.3, Appendix 3). The MMRM model including all available data up to Week 16 will be used (Section 5.2.2.2). This analysis is to support the secondary analysis of estimating the treatment difference of etrasimod 2 mg QD versus placebo for percent change from baseline in EASI at Week 16.

6.2.4. Other Exploratory Binary Endpoints

For Part 1 OLE, binary exploratory endpoints at all visits from Week 20 up to Week 68 (eg, EASI-75 response), will be analyzed at each visit, including all available data on the FAS (since intercurrent events are not considered in this Part 1 OLE phase), using the normal approximation for the difference in binomial proportions adjusting for the actual stratification factor of baseline IGA (3 vs 4) via the CMH approach (Section 5.2.1.1). The baseline IGA from the clinical database will be used as the actual stratification factor (Section 3.4.1). This CMH approach will provide treatment comparison of Etrasimod 2 mg QD versus Placebo → Etrasimod 2 mg QD and generate 95% CI for the EASI-75 response rate difference at this visit along with the 2-sided p-value (Section 5.2.1.1). Missing values will be handled by setting the EASI-75 response value to non-response, MR=NR (Section 5.3).

This CMH approach will also be applied to IGA of clear (0) or almost clear (1) response, ≥4-point reduction from baseline in PP-NRS score (analyzed only for participants with baseline PP-NRS ≥4), ≥4-point reduction from baseline in Skin Pain NRS score (analyzed only for participants with baseline Skin Pain NRS ≥4), and EASI-90 response, at all visits from Week 20 up to Week 68.

For Part 2 OL Safety, binary exploratory endpoints at all visits will only be descriptively summarized for the single treatment group of Etrasimod 2 mg QD using only observed values on the FAS (since intercurrent events are not considered in this Part 2 OL Safety phase). Therefore, missing responses will not be imputed using the MR=NR in this Part 2 OL Safety phase (Section 5.3). This approach is applicable to EASI-75 response, IGA of clear (0) or almost clear (1) response, ≥4-point reduction from baseline in PP-NRS score (analyzed only for participants with baseline PP-NRS ≥4), ≥4-point reduction from baseline in Skin Pain NRS score (analyzed only for participants with baseline Skin Pain NRS ≥4), and EASI-90 response, at all visits from Week 4 to Week 52. The baseline is the pre-dose Day 1 of Part 2 OL Safety phase (Section 3.4).

6.2.5. Other Exploratory Continuous Endpoints (Repeated Visits)

For Part 1, continuous exploratory endpoints will be descriptively summarized by treatment group and by visits for percentage change from baseline in EASI score and change from baseline in POEM.

For Part 1 OLE, continuous exploratory endpoints at all visits from Week 20 up to Week 68 (eg, percent change from baseline in EASI score), will be analyzed using a single MMRM without imputation for missing data (Section 5.2.2.2). The model will include all data from Part 1 DB Week 4 up to Part 1 OLE Week 68 on the FAS (Appendix 3). The actual stratification factor (ie, baseline IGA of 3 or 4) from clinical database will be used in the MMRM analysis. This analysis will provide the LSM of the treatment groups by visit, the LSM of the treatment difference by visit (Etrasimod 2 mg QD versus Placebo → Etrasimod 2 mg QD from Week 20 to Week 68; results from Week 4 to Week 16 will not be displayed from this analysis since they will come from a repeated analysis of the Part 1 DB phase for this endpoint) in percent change from baseline in EASI score, 95% CI and the 2-sided p-value by visit (Section 5.2.2.2). This approach is also similary applied to change from baseline in

POEM including all data at Part 1 DB Week 16 (Etrasimod 2 mg QD versus Placebo), Part 1 OLE Weeks 32 and 68 (Etrasimod 2 mg QD versus Placebo → Etrasimod 2 mg QD) on the FAS. Results from all visits will be displayed.

For Part 2 OL Safety phase, continuous exploratory endpoints at all visits will only be descriptively summarized for the single treatment group of Etrasimod 2 mg QD using only observed values on the FAS (since intercurrent events are not considered in this Part 2 OL Safety phase). Therefore, missing values will not be imputed in this Part 2 OL Safety phase. This approach is applicable to percent change from baseline in EASI score and change from baseline in POEM score, at all visits from Week 4 to Week 52. The baseline is the pre-dose Day 1 of Part 2 OL Safety phase (Section 3.4).

6.3. Subset (Subgroup) Analyses

For Part 1 DB, subgroup analyses are only performed for the primary endpoint of IGA response at Week 16 under the primary estimand (ie, Estimand E1) including only the pre-intercurrent-event data on the FAS analyzed for participants with baseline IGA ≥2. The treatment comparison of Etrasimod 2 mg QD vs. Placebo for the primary endpoint at Week 16 will be assessed at each level of the subgroup variables.

The following subgroup variables will be considered:

- Baseline IGA score – 2 levels (3, 4), based on actual clinical data.
- Sex – 2 levels (Female, Male).
- Age at baseline – 2 levels (18 - <50 years, ≥50 years).
- Race – 2 levels (White, Non-White).
- BMI at baseline – 2 levels (<30, ≥30 kg/m^2).
- AD disease duration since diagnosis by Hanifin and Rajka criteria (Hanifin and Rajka, 1980) – 2 levels (<5, ≥5 years).
- Prior systemic corticosteroid failure only – 2 levels (Yes, No).

Estimates of the treatment difference between etrasimod 2 mg QD and placebo for the primary endpoint (primary estimand) along with the 95% CI without p-value, will be presented for each level of a subgroup variable. The normal approximation using CMH approach adjusting for actual stratification factor of IGA score (Section 3.4.1) will be used for each level of a subgroup variable, except for the subgroup of baseline IGA score.

For the subgroup of baseline IGA score (2 levels of 3 or 4), the normal approximation approach to the difference in binomial proportions on the FAS will be used. For each level of baseline IGA score (3 or 4), the 95% CI of the treatment difference is calculated as:

$$(\hat{p}_A - \hat{p}_B) \pm z_{0.975}\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}$$

where $\hat{p}_A$ and $\hat{p}_B$ represent the estimated response rates for etrasimod 2 mg QD and placebo respectively, and n_A and n_B are their respective sample sizes. $z_{0.975}$ is the 97.5th percentile of the standard normal distribution. No p-value will be produced for subgroup analyses. The method of handling 0 or 100% response rate in either or both treatment groups can be found in Section 5.2.1.1.

The primary purpose of subgroup analyses is to check for consistency of the primary endpoint results across subgroup levels, ie, making sure overall results are not driven by some subset of participants (ie, particular level of a subgroup variable). Graphical display (eg, forest plot) of the treatment differences in IGA response at Week 16 (primary estimand) between etrasimod 2 mg QD and placebo by subgroups will be presented. There is no intention to have any specific inference within subgroup variables, therefore only 95% CI's are produced without corresponding p-values.

6.4. Baseline and Other Summaries and Analyses

6.4.1. Baseline Summaries

For Part 1 and Part 2 of this study, similar demographics and baseline characteristics will be included but may not be limited to the ones listed below. These characteristics will be summarized descriptively. For continuous variables, the summary will include N (which is the number of participants evaluable for the baseline characteristics), mean, median, SD and range (ie, minimum and maximum). For binary and categorical variables, the summary will include frequencies and percentages. A missing category will be included for those participants with missing value. Some ordinal variables may also be summarized as continuous variables. In addition to displays by treatment group, the summaries will also be provided for all the treatment groups combined.

Demographic characteristics:

- Baseline age (continuous in years);
- Baseline age categories (≥16 to <18, ≥18 to <50, ≥50 to <70, ≥70 years);
- Gender (Female, Male);
- Race (categories as reported in Demographics CRF);

- Ethnicity (Hispanic/Latino or Spanish origin, Not Hispanic/Latino or Spanish original, Not Reported);
- Baseline body weight (continuous in kg);
- Baseline body weight categories (<60, ≥60 to ≤100, >100 kg);
- Baseline height (continuous in cm);
- Baseline BMI (continuous kg/m^2);
- Baseline BMI categories (<18.5, ≥18.5 to <25, ≥25 to <30, ≥30 to <40, and ≥40 kg/m^2).

Baseline disease characteristics:

- AD disease duration since diagnosis by Hanifin and Rajka criteria (continuous in years);
- AD disease duration since diagnosis categories by Hanifan and Rajka criteria (<5, ≥5 years);
- Baseline IGA (3, 4) – Actual in Database (note that baseline IGA of values ≤2 will also be summarized if recorded, as they are protocol deviations);
- Baseline IGA (3, 4) – Reported in Randomization;
- Baseline EASI score (continuous);
- Baseline Total BSA (continuous in %);
- Baseline PP-NRS (continuous);
- Baseline PP-NRS categories (<4, ≥4);
- Baseline Skin Pain NRS (continuous);
- Baseline Skin Pain NRS categories (<4, ≥4);
- Baseline POEM (continuous);
- Baseline POEM categories (<4, ≥4);
- Prior Systemic Therapies (Biologics, JAK inhibitors, conventional systemic therapies).

6.4.2. Study Conduct and Participant Disposition

Study conduct and participant disposition is summarized by frequency and percentages for the following sets by treatment group for Part 1 and Part 2.

- Total number of participants screened;
- Number of participants who failed screening;
- Number of participants randomized;
- Number of participants randomized and treated;
- Number of participants in the FAS;
- Number of participants in the FAS who discontinued from the study due to any reason (for Part 1 DB phase, this will be DB phase);
- Number of participants in the FAS who completed the study (for Part 1 DB phase, this will be DB phase);
- Number of participants in the SAS.

Participants who are treated but are not randomized will also be summarized if at least one participant falls into this category.

6.4.3. Study Treatment Duration

Within any study phases (Part 1 DB, Part 1 DB+OLE, or Part 2 OL Safety), study treatment duration (in day) is defined as the time from the day of first dose (ie, Day 1) to the day of last dose of the study treatment. Besides descriptive summaries of the study treatment duration, the number of participants (%) will also be summarized by the following categories (where the upper limits correspond to the target day of the visit windows, Appendix 2.1), and ≥ categories, which are aligned with the protocol's study visits.

Table 5. Reporting Categories for Study Treatment Duration (in Day)

Part 1 DB		Part 1 DB + OLE		Part 2 OL Safety	
Categories (day)	≥ Categories (day)	Categories (day)	≥ Categories (day)	Categories (day)	≥ Categories (day)
1	≥ 1	1	≥ 1	1	≥ 1
2 – 29	≥ 2	2 – 29	≥ 2	2 – 29	≥ 2
30 – 57	≥ 30	30 – 57	≥ 30	30 – 57	≥ 30
58 – 85	≥ 58	58 – 85	≥ 58	58 – 113	≥ 58
86 - 113	≥ 86	86 – 113	≥ 86	114 – 169	≥ 114
≥ 114	≥ 114	114 – 141	≥ 114	170 – 253	≥ 170
		142 – 169	≥ 142	254 – 309	≥ 254
		170 – 225	≥ 170	310 – 365	≥ 310
		226 – 281	≥ 226	≥ 366	≥ 366
		282 – 365	≥ 282		

Table 5. Reporting Categories for Study Treatment Duration (in Day)

Part 1 DB		Part 1 DB + OLE		Part 2 OL Safety	
Categories (day)	≥ Categories (day)	Categories (day)	≥ Categories (day)	Categories (day)	≥ Categories (day)
		366 – 477	≥ 366		
		≥ 478	≥ 478		

The upper limit of period category corresponds to visit window's target day (target week x 7 days + 1 day).

6.4.4. Prior and Concomitant Medications and Nondrug Treatments

Prior medications and non-drug treatments which start prior to baseline (ie, Day 1) will be summarized by treatment group and study phases. Concomitant medications and non-drug treatments which were taken any time during the study treatment period will also be summarized by treatment group and study phases. In addition, Concomitant (Day 1) medication and non-drug treatments may also be summarized for those medications and non-drug treatments taken at baseline (ie, Day 1). All summaries will be coded to the WHODrug version at the time of database release and this version will be noted in the summaries.

6.5. Safety Summaries and Analyses

All the safety data will be summarized descriptively through appropriate data tabulations, descriptive statistics, and graphical presentations. Treatment-emergent AEs including SAEs and AE leading to discontinuation of study treatment, vital signs, 12-lead ECG parameters, and laboratory tests, etc. will be summarized according to the Clinical Data Interchange Standards Consortium (CDISC) and Pfizer Standards (CaPS). All safety analyses will be performed on the SAS (Section 4.2).

6.5.1. Adverse Events

The MedDRA version for AE coding at the time of database release will be adopted and the version used will be displayed in all relevant safety data tabulations. TEAEs will be displayed by these 3 phases: (1) Part 1 DB (up to Week 16), (2) Part 1 DB+OLE (up to Week 68), (3) Part 2 (up to Week 52), using the same treatment group labels (Section 5.2). Safety data from follow-up visits will also be be included for summaries.

In general, all treatment-emergent AE are tabulated in standard safety displays. AESIs will be summarized descriptively only (Section 3.5.2). The number and percent of participants reporting specific past and present medical histories at screening or baseline will be summarized by medical history reporting term (MedDRA) and treatment group following CaPS.

6.5.2. Laboratory Data

All relevant laboratory tests will be summarized. These summaries include categorical tables (eg, normal, high, low), and descriptive statistics for change or percent change from baseline by treatment group and visit. These laboratory tests may include the following but are not limited to this list: hematology (including TBNK and coagulation Panel), blood chemistry,

urinalysis, liver function assessment, etc. Participants who meet clinically significant categories or defined abnormal ranges as defined by laboratory will also be summarizd.

6.5.3. Vital Signs

Absolute values and changes from baseline in systolic and diastolic BP and pulse rate will be summarized by treatment group and time post-dose in accordance with the CaPS. The number and percent of participants with clinically significant change from baseline will be tabulated by treatment group.

6.5.4. Electrocardiograms

Changes from baseline for the ECG parameters QT interval, heart rate, QTc interval, PR interval, and QRS complex will be summarized by treatment and time.

Changes from baseline for the ECG parameters HR, QTcF, PR interval, and QRS complex will be summarized by treatment and time. The frequency of uncorrected QT values above 500 ms will be tabulated.

The number (%) of participants with maximum post-dose QTcF values and maximum increases from baseline in the following categories will be tabulated by treatment.

Table 6. Safety QTcF Assessment

Degree of Prolongation	Mild (ms)	Moderate (ms)	Severe (ms)
Absolute value	>450 – 480	>480 – 500	>500
Increase from baseline		30-60	>60

6.5.5. Physical Examination

This is not applicable as physcial examination is not collected in the study database.

6.5.6. Protocol Deviations and COVID-19 Related Reporting

In order to describe the impact of COVID-19, participants who discontinue from the study due to COVID-19 as the primary reason will be summarized in the disposition table (Section 6.4.2). The specific reasons for participants who discontinue from the study due to COVID-19, if collected in detail, such as participant decision, participant self-isolation, non-compliance with protocol, transportation issue, site restrictions, physician decision, site closed, lack of drug at site, sponsor decision, and severity of COVID-19 infection will be presented in the disposition listing.

The number and percentage of participants with at least one important PD and with at least one PD associated with COVID-19 will be summarized for the Part 1 DB phase by treatment group, deviation category, and deviation sub-category using the FAS. Important PDs and PDs related to COVID-19 will be summarized similarly overall using the FAS. This is also performed similarly for Part 1 OLE phase and Part 2 OL Safety phase.

A listing of all PDs will be provided. Important PDs and COVID-19 related PDs will be flagged in this listing for the Part 1 DB phase, Part 1 OLE phase, and Part 2 OL Safety phase.

7. INTERIM ANALYSES

An interim analysis (IA) will be conducted by the Sponsor in an unblinded manner to assess efficacy and safety after all applicable Part 1 participants complete or discontinue from study prior to the 16-week DB treatment period. The IA results will be used for decisions regarding stopping for futility and internal program development. Study conduct, including Part 1 participant continuation in OLE, will continue while the data are prepared and analyzed for the IA. The Part 2 enrollment will not be triggered until after IA shows an acceptable safety and efficacy of etrasimod 2 mg QD in the target population. Based on the IA results, the study may continue, or the study may be terminated early.

Specifically, the scope of this IA includes the following. When all applicable Part 1 participants have completed the 16-week DB treatment phase, an interim datacut will be performed that will include all efficacy and safety data up to Week 16 visit. After data cleaning is performed, unblinding of Part 1 DB treatment assignments will take place, and interim database will be released for the IA to be performed by the Sponsor and the study team. Investigators and other site staff will remain blinded during the OLE to participants' treatment assignment received in the Part 1 DB phase until the end of Part 1 OLE. Efficacy data summary will include the assessment of primary efficacy endpoint, secondary efficacy endpoints, and exploratory efficacy endpoints for Part 1 DB (Section 2.2) using the pre-specified statistical methodologies (Section 5) to perform all relevant analyses described for Part 1 DB including subgroup analysis (Section 6.3). Safety data summary will include the assessment of all relevant safety endpoints (Section 3.5) descriptively summarized up to end of Part 1 DB phase.

Prior to this formal Part 1 DB IA, Sponsor may perform earlier unblinded interim analyses to inform internal decision-making regarding the development program. These earlier interim analyses will be conducted and reviewed by a limited number of team members independent of the study team. Details on these earlier unblinded interim analyses will be described in an unblinding operational plan.

When Part 1 OLE treatment phase is completed, the final analysis will be performed based on the final released official database for Part 1 of this study. This final analysis of Part 1 will include repeating the IA for Part 1 DB phase using this final data, as well as cumulative analyses including the final data from the Part 1 OLE phase. There will be no IA planned for Part 2 OL Safety phase of this study. When Part 2 OL Safety treatment phase is completed, the final analysis will be performed based on the final released official database for Part 2 of this study.

The e-DMC will be responsible for ongoing monitoring of the safety of participants in the study according to the charter. The recommendations made by the e-DMC to alter the conduct of the study will be forwarded to Pfizer for final decision. Information about the e-DMC can be found in the e-DMC Charter, which outlines the operating procedures of the

committee, including specific description of the scope of their responsibilities, including a plan where communication timelines are defined.

8. REFERENCES

1. Cochran WG. (1954). Some methods of strengthening the common X^2 tests. Biometrics. 10: 417-451.

2. Mantel N, Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. J. Natl. Cancer Inst. 22:719-748.

3. Agresti A. (2002). Categorical Data Analysis, 2nd Edition, New York: John Wiley & Sons.

4. Miettinen O, Nurminen M. (1985). Comparative analysis of two rates. Statistics in Medicine. 4: 213-226.

5. Hanifin, JM, Rajka, G. (1980). Diagnostic features of atopic dermatitis. Acta Dermatovener (Stockholm). Suppl 92:44-47.

6. Carpenter JR, Roger JH, Kenward MG. (2013). Analysis of longitudinal trials with protocol deviation: a framework for relevant, accessible assumptions, and inference via multiple imputation. Journal of Pharmaceutical Statistics. 23: 1352-1371.

7. Rubin DB. (1987). Multiple imputation for nonresponse in surveys. New York: Wiley.

9. APPENDICES

Appendix 1. Summary of Efficacy Analyses

Table 7. Table of Efficacy Analyses

Endpoint Type	Estimand	FAS	Part 1 DB Phase	Part 1 DB + OLE Phase	Part 2 OL Safety
Binary Endpoints	**Estimand E1 (Pre-Intercurrent Event Data)**	**FAS for Estimand E1**	Method: CMH method adjusting to actual stratification factor of IGA score (3 or 4), MR=NR • IGA Response at Week 16 (Primary Endpoint), participants with baseline IGA ≥2 • EASI-75 Response at Week 16 (Secondary Endpoint) • IGA Response at all timepoints except Week 16, participants with baseline IGA ≥2 • EASI-75 Response at all timepoints except Week 16 • ≥4-point reduction from baseline in PP-NRS at all timepoint, participants with baseline PP-NRS ≥4 • ≥4-point reduction from baseline in Skin Pain NRS at all timepoints, participants with baseline Skin Pain NRS ≥4 • EASI-90 Repsonse at all timepoints • ≥4-point reduction from baseline in POEM at Week 16, participants with baseline POEM ≥4	The analyses for Part 1 DB Phase will be repeated using Part 1 final database.	None

Table 7. Table of Efficacy Analyses

Endpoint Type	Estimand	FAS	Part 1 DB Phase	Part 1 DB + OLE Phase	Part 2 OL Safety
	Estimand E1a (All Data)	**FAS for Estimand E1a**	Method: CMH method adjusting to actual stratification factor of IGA score (3 or 4), MR=NR • IGA Response at Week 16 (Primary Endpoint), participants with baseline IGA ≥2 • IGA Response at all timepoints except Week 16, participants with baseline IGA ≥2 • EASI-75 Response at Week 16 (Secondary Endpoint) • EASI-75 Response at all timepoints except Week 16	The analyses for Part 1 DB Phase will be repeated using Part 1 final database.	None
	No Estimand (All data)	**FAS**	None	Method: CMH method adjusting to actual stratification factor of IGA score (3 or 4), MR=NR • EASI-75 Response at all timepoints • IGA of clear (0) or almost clear (1) based on 5-point scale at all timepoints • ≥4-point reduction from baseline in PP-NRS at all timepoints, participants with baseline PP-NRS ≥4 • ≥4-point reduction from baseline in Skin Pain NRS at all timepoints, participants with baseline Skin Pain NRS ≥4 • EASI-90 Repsonse at all timepoints	Method: Descriptive summary using observed data • EASI-75 Response at all timepoints • IGA of clear (0) or almost clear (1) based on 5-point scale at all timepoints • ≥4-point reduction from baseline in PP-NRS at all timepoints. participants with baseline PP-NRS ≥4 • ≥4-point reduction from baseline in Skin Pain NRS at all timepoints, participants with baseline Skin Pain NRS ≥4 • EASI-90 Repsonse at all timepoints

Table 7. Table of Efficacy Analyses

Endpoint Type	Estimand	FAS	Part 1 DB Phase	Part 1 DB + OLE Phase	Part 2 OL Safety
Continuous Endpoints	Estimand E2 (Pre-Intercurrent Event Data)	FAS for Estimand E2	Method: JTC method via multiple imputations of missing data including data up to Week 16 • Percent change from baseline in EASI score at Week 16 (Secondary Endpoint) • Percent change from baseline in EASI score at all timepoints except Week 16	The analyses for Part 1 DB Phase will be repeated using Part 1 final database.	None
	Estimand E2a (All Data)	FAS for Estimand E2a	Method: MMRM without imputation of missing data including data up to Week 16 • Percent change from baseline in EASI score at Week 16 (Secondary Endpoint) • Percent change from baseline in EASI score at all timepoints except Week 16	The analyses for Part 1 DB Phase will be repeated using Part 1 final database.	
	No Estimand (All data)	FAS	None	Method: MMRM without imputation of missing data including data at all timepoints (Report only OLE visits for Percent change from baseline in EASI score, report all visits for Change from baseline in POEM) • Percent change from baseline in EASI score at all timepoints • Change from baseline in POEM at Weeks 16, 32 and 68	Method: Descriptive summaries using observed data • Percent change from baseline in EASI score at all timepoints • Change from baseline in POEM at Weeks 24 and 52

Appendix 2. Data Derivation Details

Appendix 2.1. Definition and Use of Analysis Visit Windows in Reporting

There are two sets of analysis visit windows, one set for Part 1 (both DB and OLE phases) and another set for Part 2 (OL). For reporting purposes, the analysis visit windows will be used for efficacy and safety display that show by scheduled visits. Observations from all visits including scheduled, unscheduled and early termination visits will all be eligible for allocation to an analysis visit. If two or more observations fall into the same visit window, the observation closest to the target day will be used in the analysis. If there is a tie, the later observation will be used. However, for a participant who discontinues from the study early prior to Week 68 visit window in Part 1 (similarly to Week 52 visit window in Part 2), if two or more observations fall in the same visit window prior to Week 68 in Part 1 (similarly to Week 52 in Part 2), then the latest observation (rather than the observation closest to the target day) within this visit window will be used in the analysis for that visit window.

The analysis visit windowing will be applied after intercurrent events (initation of prohibited medications or discontinuation of study treatment, whichever is earlier) have been accounted for, that is, to the pre-intercurrent-event data and to all data separately (Appendix 3).

Table 8. Analysis Visit Windows for Part 1

Visit Label	Target Day[a]	Window Definition
Baseline	Day 1	Last non-missing assessment on or before Day 1 and prior to first dose of investigational product (ie, etrasimod or placebo)
Week 4	Day 29	Day 2 to Day 42
Week 8	Day 57	Day 43 to Day 70
Week 12	Day 85	Day 71 to Day 98
Week 16[b]	Day 113	Day 99 to [Treatment Switch Day[c] or Day 126, whichever is smaller]
Week 20	Day 141	[Treatment Switch Day[c]+1 or Day 127, whichever is smaller] to Day 154
Week 24	Day 169	Day 155 to Day 196
Week 32	Day 225	Day 197 to Day 252
Week 40	Day 281	Day 253 to Day 322
Week 52	Day 365	Day 323 to Day 420
Week 68	Day 477	≥ Day 421

a. Target day is equal to the target week×7 days + 1 day.
b. All data collected on Week 16 visit, first open-label dose day (treatment switch day), will be included in ≤ Week 16 except for AE data, which will be included in ≥ Week 20. Any new treatment-emergent AE occurred on the day of first open-label dose (treatment switch day) will be included in ≥ Week 20.
c. Day of first open-label dose at Week 16 = Treatment switch day (ie, treatment switch day = treatment switch date – date of the first dose +1), where treatment switch date is taken from the Treatment Dosing CRF, as the start date in the first entry of the Treatment Dosing log for the period of Week 16 to Week 20.

Note that the upper limit of the Week 16 visit window and the lower limit of the Week 20 visit window are defined uniquely for each of the participants in the study regardless of which treatment group the participants are randomized to.

If treatment switch day (ie, day of first open-label etrasimod dose) falls between Day 99 and Day 126, then the upper limit of the Week 16 visit window is the smaller of treatment switch day or Day 126, and the lower limit of the Week 20 visit window is the smaller of treatment switch day+1 or Day 127.

In the unlikely case if treatment switch day < Day 99, > Day126, or missing (eg, discontinuation of study treatment or study prior to Week 16), then the upper limit of the Week 16 visit window will be set to Day 126 and the lower limit of the Week 20 visit window will be set to Day 127.

Table 9. Analysis Visit Windows for Part 2

Visit Label	Target Day[a]	Window Definition
Baseline	Day 1	Last non-missing assessment on or before Day 1 and prior to first dose of etrasimod
Week 4	Day 29	Day 2 to Day 42
Week 8	Day 57	Day 43 to Day 84
Week 16	Day 113	Day 85 to Day 140
Week 24	Day 169	Day 141 to Day 210
Week 36	Day 253	Day 211 to Day 280
Week 44	Day 309	Day 281 to Day 336
Week 52	Day 365	≥ Day 337

a. Target day is equal to the target week×7 days + 1 day.

For safety data collected for participants who complete the 2-week and 4-week follow-up visits (Weeks 70 and 72 in Part 1 OLE, Weeks 54 and 56 in Part 2 OL Safety phase), available safety data will be summarized for these nominal visits.

For 12-Lead ECG collected on Day 1 during first-dose cardiac monitoring, the labels for timepoints can be “Day 1 Predose”, “Day 1 1HR Postdose” to “Day 1 4HR Postdose”.

Appendix 2.2. Endpoint Derivations

Appendix 2.2.1. Investigator’s Global Assessment (IGA)

IGA of AD is scored on a 5-point scale (0-4), reflecting a global consideration of the erythema, induration, and scaling. The clinical evaluator of AD will perform an assessment of the overall severity of AD and assign an IGA score: 0=clear, 1=almost clear, 2=mild, 3=moderate, or 4=severe. The assessment will be a static evaluation without regard to the

score at a previous visit. According to protocol's inclusion criteria, only participants with IGA score ≥3 at screening and baseline (Day 1) are included in this study. The following binary efficacy endpoints will be derived based on IGA score at a visit.

- IGA response: IGA score of clear (0) or almost clear (1) on a 5-point scale and a reduction from of ≥2 from baseline. Only participants with baseline IGA ≥2 are included in all analyses of this endpoint.

- IGA score of clear (0) or almost clear (1) on a 5-point scale.

Appendix 2.2.2. Body Surface Area (BSA)

BSA affected by AD will be determined by the participant's handprint method, where the full hand of the participant (ie, the participant's fully extended palm, fingers, and thumb together in a closed position) represents approximate 1% of the total BSA. The BSA AD involvement ranges from 0% to 100%, with higher values representing greater severity of AD. The BSA is performed separately for 4 regions of the body: Region 1 – head and neck; Region 2 – trunk (including genital area); Region 3 – upper limbs; and Region 4 – lower limbs (including buttocks). The total number of handprints in each body region is given in the following table. The total number of handprints summed up across the 4 body regions is 100, which is used as the denominator for total BSA. Total BSA efficacy (%) will be derived from the sum of the BSA in handprints across 4 body regions (out of total of 100), and it is assessed as part of the EASI assessment (Appendix 2.2.3). According to protocol's inclusion criteria, only participants with BSA ≥10% of AD involvement at screening and baseline (Day 1) are included in this study.

Table 10. Handprint Determination of BSA

Body Region	Total Number of Handprints in Body Region*	Surface Area of Body Region Equivalent of One Handprint*
Head and Neck (h)	10	10%
Upper Limbs (u)	20	5%
Trunk (including axillae and groin/genitals) (t)	30	3.33%
Lower Limbs (including buttocks) (l)	40	2.5%

*Handprint refers to the hand size of each individual participant.

Further, the extent (%) of each of the four body region (h, u, t, and l) will be determined by dividing the number of BSA handprints by the total (maximum) number of handprints for that body region (range: 0% to 100%). For example, if there are 3 BSA handprints for Head and Neck region, the extent (%) of the Head and Neck region is (3/10) x 100% = 30%. This extent (%) to which each of the four body regions (h, u, t, and l) is involved with atopic dermatitis is categorized (ie, mapped) to a numerical Area Score (Ah, Au, At, and Al, respectively) using a non-linear scaling method according to the following BSA scoring criteria. Using the same previous example, Ah is 3.

Table 11. Mapping of Body Region BSA Percentage to Area Score

Percent BSA with AD in a body region (h, u, t, and l)	Area Score (Ah, Au, At, and Al)
0%	0
>0 - <10%	1
10 - <30%	2
30 - <50%	3
50 - <70%	4
70 - <90%	5
90 - 100%	6

These Area Scores (Ah, Au, At, and Al) of the four body regions will be used to derive the EASI score (Appendix 2.2.3). If any one of the four body regions has missing number of BSA handprints, the extent (%) of BSA in this body region will be missing and so will the Area Score for this body region. The total BSA (%) will also be missing.

Appendix 2.2.3. Eczema Area and Severity Index (EASI)

The EASI quantifies the severity of a participant's AD based on both severity of lesion clinical signs and the percent of BSA affected. EASI is a composite scoring by the atopic dermatitis clinical evaluator of the degree of erythema, induration/papulation, excoriation, and lichenification (each scored separately) for each of four body regions, with adjustment for the percent of BSA involved for each body region and for the proportion of the body region to the whole body.

There are 4 lesion clinical signs for each of the 4 body regions (Appendix 2.2.2): erythema (E), induration/papulation (I), excoriation (Ex) and lichenification (L). For each of the lesion clinical signs, severity is scored on a 4-point scale: 0=absent, 1=mild, 2=moderate, and 3=severe. These 16 scores (0-3) are denoted as follows. Morphologic descriptors for each clinical sign severity score can be found in the study protocol.

Table 12. Symbols Denoting Lesion Clinical Signs by Body Region

Body Regions (row)/Clinical Signs (column)	Erythema (E)	Induration/Papulation (I)	Excoriation (Ex)	Lichenification (L)
Head and Neck (h)	Eh	Ih	Exh	Lh
Upper Limbs (u)	Eu	Iu	Exu	Lu
Trunk (including axillae and groin/genitals) (t)	Et	It	Ext	Lt
Lower Limbs (including buttocks) (l)	El	Il	Exl	Ll

In EASI scoring, each body region carries different body region weighting, the sum of these weightings is equal to 1: Head and Neck (0.1), Upper Limbs (0.2), Trunk (including axillae and groin/genitals) (0.3) and Lower Limbs (including buttocks) (0.4).

In each body region, the sum of the Clinical Signs Severity Scores for erythema, induration/papulation, excoriation, and lichenification is multiplied by the Area Score (Appendix 2.2.2) and by the body region weighting to provide a body region value (ie, body region subscore), which is then summed across all four body regions resulting in an EASI score as described in the following equation.

EASI = 0.1 Ah (Eh + Ih + Exh + Lh) + 0.2 Au (Eu + Iu + Exu + Lu) + 0.3 At (Et + It + Ext + Lt) + 0.4 Al (El + Il + Exl + Ll)

The EASI score can vary in increments of 0.1 and range from 0.0 to 72.0, with higher scores representing greater severity of AD. According to protocol's inclusion criteria, only participants with EASI ≥16 at screening and baseline (Day 1) are included in this study. The following efficacy endpoints will be derived based on the EASI scores assessed at baseline and at post-baseline visits:

- Percent change from baseline in EASI score, defined as the change from baseline in EASI score divided by baseline EASI score, scaled to %.
- EASI-50, defined as a response if ≥50% reduction from baseline in EASI score (as a pre-specified discontinuation criterion rather than an efficacy endpoint).
- EASI-75, defined as a response if ≥75% reduction from baseline in EASI score.
- EASI-90, defined as a response if ≥90% reduction from baseline in EASI score.

If any component score (BSA Area Score or lesion clinical sign severity score) in the equation is missing at a visit, the EASI will be missing at that visit. The derived endpoints will also be missing at that visit.

Appendix 2.2.4. Peak Pruritus Numerical Rating Scale (PP-NRS)

PP-NRS is a single-item on the scale from 0 (no itch) to 10 (worst itch imaginable), NRS rating scale, to assess itch at the worst moment during the previous 24 hours. The following binary efficacy endpoint will be derived: ≥4-point reduction from baseline in PP-NRS. Only participants with baseline PP-NRS ≥4 are included in all analyses of this endpoint. If PP-NRS score is missing for a visit, the derived endpoint will also be missing at that visit.

Appendix 2.2.5. Skin Pain Numerical Rating Scale (Skin Pain NRS)

The Skin Pain NRS is a single-item on the scale from 0 (no skin pain) to 10 (worst skin pain imaginable), NRS rating scale, to assess the skin pain at its worst over the previous 24 ours. The following binary efficacy endpoint will be derived: ≥4-point reduction from baseline in Skin Pain NRS. Only participants with baseline Skin Pain NRS ≥4 are included in all analyses of this endpoint. If Skin Pain NRS score is missing for a visit, the derived endpoint will also be missing at that visit.

Appendix 2.2.6. Patient-Oriented Eczema Measure (POEM)

POEM is a 7-item measurement used for monitoring symptom severity of atopic eczema. To address severity, the POEM measures 7 aspects of the experiences of participants with eczema including skin itch, sleep disturbance, skin bleeding, weeping/oozing of the skin, cracked skin, flaking skin, and skin dryness/roughness. Participants are asked to answer based on the number of days they experienced each on a 5-point scale (0=No days, 1=1 to 2 days, 2=3 to 4 days, 3=5 to 6 days, and 4=everyday) over the past week based on recall. The scores from the 7 questions are accumulated to provide an overall POEM score with a range of 0 to 28. An overall POEM score is as follows: 0 to 2 = 'clear/almost clear'; 3 to 7 ='mild'; 8 to 16 = 'moderate'; 17 to 24 = 'severe'; and 25 to 28 = 'very severe' atopic eczema.

POEM is measured at baseline, Weeks 16, 32, and 68 in Part 1; baseline, Weeks 24 and 52 in Part 2. The following efficacy endpoints will be derived using POEM:

- Change from baseline in POEM.
- ≥4-point reduction from baseline in POEM. Only participants with baseline POEM ≥4 are included in all analyses of this endpoint.

If POEM score is missing at a visit, the derived endpoints will also be missing at that visit.

Appendix 2.2.7. Range of Values for Continuous Efficacy Endpoints

The following table displays the unit, theoretical range of values, and the direction of improvement for each of select continuous efficacy endpoints.

Table 13. Numerical Characteristics of Select Continuous Scores and Derived Effiacy Endpoints

Endpoint	Unit	Theoretical Range of Values	Direction of Improvement from Baseline
IGA	None	0-4	Decrease from Baseline
Total BSA	%	0-100	Decrease from Baseline
EASI	None	0-72	Decrease from Baseline
PP-NRS	None	0-10 (NRS)	Decrease from Baseline
Skin Pain NRS	None	0-10 (NRS)	Decrease from Baseline
POEM	None	0-28	Decrease from Baseline

Appendix 3. Efficacy Data Set Descriptions For Part 1 Double-Blind Phase

There are two intercurrent events considered for Part 1 DB phase of this study, (1) the initiation of prohibited medications, and (2) the discontinuation of the study treatment (etrasimod 2 mg QD or placebo) due to any reason.

During the Part 1 DB phase, if a participant initiates prohibited medications (intercurrent event 1 above) and/or discontinues study treatment due to any reason (intercurrent event 2 above), the date of the earlier intercurrent event is noted (as recorded on relevant CRF page) and is used to differentiate and to define the data up to Week 16 into "Pre-intercurrent-event data", "Post-intercurrent-event data", and "All data" (ie, both Pre- and Post-intercurrent-event data together), as laid out in the table below. Only efficacy data are in scope for this data differentiation.

Table 14. Efficacy Data Set Differentiation by Estimand in Part 1 DB Phase

Estimand	Participants who initiate prohibited medications and/or discontinue study treatment due to any reason prior to Week 16	Participants who do not experience any intercurrent events
E1	For this participant, the date/day of the earlier intercurrent event is flagged and noted (eg, Day X relative to baseline Day 1). Only data collected/assessed on or before Day X+2 (ie, allowing up to 2 additional days after date of earlier intercurrent event), as pre-intercurrent-event data, are included in this dataset. Any data collected/assessed after Day X+2, as post-intercurrent-event data, are excluded in this dataset. Visit windowing is performed after this data differentiation on the Pre-intercurrent-event data.	All data (unaffected due to no intercurrent events) of these participants will be included.
E1a	All data of these participants will be included regardless of the intercurrent events.	Same as E1 above.
E2	Same as E1 above.	Same as E1 above.
E2a	Same as E1a above.	Same as E1a/E1 above.
No Estimand	All data of these participants and available data will be included regardless of the intercurrent events.	

All efficacy data collected in Part 1 OLE phase and Part 2 OL Safety phase will include all data collected/assessed regardless of these intercurrent events (ie, no estimand defined). All safety data will be analyzed regardless of these intercurrent events.

Appendix 4. Initiation of Prohibited Medications

The intercurrent event of initiation of prohibited medications will be identified using the final reconciliated PIPD dataset. The PD category of "concomitatnt medications" and the subcategory of "took prohibited concomitant medication/vaccine – meeting intercurrent event" will be used to identify this protocol deviation as the intercurrent event. This protocol deviation has to be verified using participant's entries of concomitant medications via clinical review. Once verified, the analysis population exclusion (APE) status will be checked "Yes" and this protocol deviation is deemed to be the intercurrent event of initiation of prohibited medications. The onset date of the protocol deviation as recorded on the PIPD dataset will

serve as the date of initiation of prohibited medications. This is applicable to the Part 1 DB phase only.

Appendix 5. List of Abbreviations

The following table provides a comprehensive list of abbreviations. Mathematical symbols are not included on this list.

Table 15. Table of Abbreviations

Abbreviation	Term
AD	atopic dermatitis
AE	adverse event
AESI	adverse event of special interest
ANCOVA	analysis of covariance
APE	analysis population exclusion
AV	atrioventricular
BMI	body mass index
BP	blood pressure
CaPS	Clinical Data Interchange Standards Consortium and Pfizer Standards
CDISC	Clinical Data Interchange Standards Consortium
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus Disease 2019
CRF	case report form
CS	compound symmetry
CSH	heterogeneous compound symmetry
CSR	clinical study report
CT SAE	clinical trial serious adverse event
DB	Double-Blind
e-DMC	external data monitoring committee
E1	Estimand E1 strategy
E1a	Estimand E1a strategy
E2	Estimand E2 strategy
E2a	Estimand E2a strategy
EASI	Eczema Area and Severity Index
EASI-50	≥50% reduction from baseline in EASI score
EASI-75	≥75% reduction from baseline in EASI score
EASI-90	≥90% reduction from baseline in EASI score
ECG	electrocardiogram
FAS	Full Analysis Set
HR	heart rate
IA	interim analysis
IGA	Investigator's Global Assessment

Table 15. Table of Abbreviations

Abbreviation	Term
JTC	Jump-to-Control
LSM	least-squares mean
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram
MMRM	mixed-effects model with repeated measures
MR=NR	missing response as non-response
ms	millisecond
NA	Not Applicable
NRS	Numerical Rating Scale
OL	Open-Label
OLE	Open-Label Extension
PACL	Protocol Administrative Change Letter
PD	Protocol Deviation
PIPD	Potentially Important Protocol Deviation
PML	Progressive multifocal leukoencephalopathy
POEM	Patient-Oriented Eczema Measure
PP-NRS	Peak Pruritus Numerical Rating Scale
PRES	posterior reversible encephalopathy syndrome
PR	pulse rate
QD	once daily
QTc	corrected QT
QTcF	corrected QT (Fridericia method)
S1PRM	Sphingosine-1-phosphate receptor modulators
SAE	serious adverse event
SAP	statistical analysis plan
SAS	Safety Analysis Set / Statistical Analysis System
SD	standard deviation
SE	standard error
TBNK	T-lymphocyte, B-Lymphocyte, and Natural Killer cell
TCS	topical corticosteroids
WHODrug	World Health Organization Drug Dictionary