STATISTICAL ANALYSIS PLAN

APD334-206 (C5041009)

A Phase 2 Randomized, Double-Blind, Placebo-Controlled Study to Assess the Safety and Efficacy of Etrasimod in Adult Subjects with Eosinophilic Esophagitis

AUTHOR: PPD

VERSION NUMBER AND DATE: V 1.0, 12Jan2023

Document : [redacted]

Author: PPD

Version Number: 1.0

Version Date: 12Jan2023

[redacted]

Reference: [redacted]

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan V1.0 (Dated 12 Jan. 2023) for Protocol APD334-206.

	Name	Signature	Date (DDMMYYYY)
Approved By:	PPD	Refer to eSignature	
Position:	PPD		
Company:	IQVIA		

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan, authorizing that the content is acceptable for the reporting of this study.

	Name	Signature	Date (DDMMYYYY)
Approved By:	PPD		
Position:	PPD		
Company:	IQVIA		
	Name	Signature	Date (DDMMYYYY)
Approved By:	PPD		
Position:	PPD		
Company:	Pfizer, Inc.		

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
0.1	19FEB2021	PPD	Not Applicable – First Version
0.2	21Jul2021	PPD	Updated based on Legacy Arena draft 1 comments and alignment with other ongoing etrasimod studies, including APD334-301 and APD334-302.
0.3	01Mar2022	PPD	Update based on Legacy Arena draft 2 comments and alignment with Version 1 SAPs for APD334-301 and APD334-302. The analysis visit window (study day) ranges are updated to be "212 to 274" for Week 32, and "275 to 344" for Week 46. Baseline eosinophil PEC (eos/hpf) is removed from section 8.2 as it is not a randomization stratification factor.
0.4	23Aug2022	PPD	"Additionally, a listing of retinal photograph and eye pressure assessments will be provided for subjects who experienced an AE related to eye disorders." is removed from 17.7.3. Kye from Pfizer confirmed the removal on 23Aug2022.
1.0	Dec2022	PPD	Mainly updates are as follows: Added the Identification of Rescue Therapy guidance in Appendix to document/align with the Rescue Therapy identification process Updated the DSQ total score calculation Updated the visit window language

			Updated the model for primary endpoint by using rank score per protocol amendment v3.0 (04Jan2023). Added/updated the safety analyses in extension period per protocol amendment v3.0 (04Jan2023).

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Explanation
ADL	activities of daily living
ADR	adverse drug reaction
AE	adverse event
AESI	adverse event of special interest
ALC	absolute lymphocyte count
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
AV	atrioventricular
AZA	azacytidine
β-hCG	beta-human chorionic gonadotropin
BP	blood pressure
bpm	beats per minute
CBC	complete blood count
CD	Crohn's disease
CFR	Code of Federal Regulations
cGMP	current Good Manufacturing Practices
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CMP	Clinical Monitoring Plan
COPD	chronic obstructive pulmonary disease
COVID-19	coronavirus disease 2019
CRA	clinical research associate
CRO	contract research organization
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CYP	cytochrome P450
DLCO	diffusing capacity of the lungs for carbon monoxide
DNA	deoxyribonucleic acid
DSQ	Dysphagia Symptom Questionnaire
ECG	electrocardiogram
eCRF	electronic case report form

Abbreviation	Explanation
eDiary	electronic diary
EEn	eosinophilic enteritis
EG	eosinophilic gastritis
EGD	esophagogastroduodenoscopy
eGFR	estimated glomerular filtration rate
EGID	eosinophilic gastrointestinal disease
EoE	eosinophilic esophagitis
EoE-HSS	Eosinophilic Esophagitis Histology Scoring System
EoE-QOL-A	Adult Eosinophilic Esophagitis Quality of Life
eos	eosinophils
EREFS	Eosinophilic Esophagitis Endoscopic Reference Score
ET	Early Termination
FA	fluorescein angiogram
FAQ	Food Avoidance Question
FCS	fully conditional specification
FDA	Food and Drug Administration
FEV 1	forced expiratory volume at 1 second
FAS	Full Analysis Set
FSH	follicle-stimulating hormone
FVC	forced vital capacity
GCP	Good Clinical Practice
GERD	gastroesophageal reflux disease
GI	gastrointestinal
GWAS	genome-wide association studies
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
hpf	high power field
HR	heart rate
HSS	Histology Scoring System
IAC	independent adjudication committee
IB	Investigator's Brochure

Abbreviation	Explanation
IC	informed consent
ICF	informed consent form
ICH	International Council for Harmonization
IEC	Independent Ethics Committee
IgE	immunoglobulin E
IL	interleukin
IMID	immune-mediated inflammatory disorder
IND	Investigational New Drug
INR	international normalized ratio
IRB	Institutional Review Board
IUD	intrauterine device
IV	intravenous
IWRS	Interactive Web Response System
JAK	Janus kinase
LS	least square
6-MP	6-mercaptopurine
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mantel-Haenszel
mFAS	modified Full Analysis Set
MI	multiple imputation
MMF	mycophenolate mofetil
MTX	methotrexate
NASH	nonalcoholic steatohepatitis
OCT	optical coherence tomography
OIT	oral immunotherapy
PD	pharmacodynamic(s)
PDMP	Protocol Deviation Management Plan
PEC	peak eosinophil count
PFT	pulmonary function test
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PHE	public health emergency

Abbreviation	Explanation
PK	pharmacokinetic(s)
PML	progressive multifocal leukoencephalopathy
PPI	proton pump inhibitor
PPI-REE	proton pump inhibitor-responsive esophageal eosinophilia
PTEN	phosphatase and tensin homolog
qd	once daily
QOL	quality of life
QTcB	Bazett's corrected QT interval
QTcF	Fridericia's corrected QT interval
RNA	ribonucleic acid
S1P	sphingosine 1-phosphate
$S1P_{1,4,5}$	sphingosine 1-phosphate receptors "1," "4," 5
SAE	serious adverse event
SAM	severe atopy metabolic wasting
SAP	Statistical Analysis Plan
SCIT	subcutaneous immunotherapy
SD	standard deviation
SLIT	sublingual immunotherapy
SOP	standard operating procedure
SpO2	arterial oxygen saturation
6-TG	6-thioguanine
TBNK	"T," "B," "NK," CD4+ T cell and CD8+ T cell count (absolute number and percentage)
TEAE	treatment-emergent adverse event
TGF-β	transforming growth factor-beta
Th2	type 2 helper T
TNFα	tumor necrosis factor alpha
TSLP	thymic stromal lymphopoietin
UC	ulcerative colitis
ULN	upper limit of normal
WBC	white blood cell
WOCBP	woman of child-bearing potential

1. INTRODUCTION

This Statistical Analysis Plan (SAP) describes the statistical rationale, methods, rules, and conventions to be used in the presentation and analysis of efficacy, safety, pharmacokinetic (PK), CCI [redacted] for Clinical Study Protocol APD334-206. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed (ICH 1998). This SAP is based on the Clinical Study Protocol APD334-206 Amendment 3.0, dated XX January 2023 and will cover the whole study period (i.e., Double-Blind Treatment Period and Extension Treatment Period). CCI [redacted]

The table, figure, and listing (TFL) shells are prepared in a separate file based on this analysis plan. Upon approval of the SAP, some updates (including, but not limited to titles, footnotes, headings and re-numbering of tables) on TFL shells are allowed without SAP amendment if the updates within TFL shells do not conflict with the contents of the SAP.

2. STUDY OBJECTIVES AND ESTIMANDS

2.1. PRIMARY OBJECTIVE

The primary objectives are:

- To evaluate the effects of etrasimod on esophageal eosinophilia in adult subjects with active eosinophilic esophagitis (EoE)
- To evaluate the dose-response relationship of 2 doses of etrasimod versus placebo in adult subjects with active EoE
- To select an etrasimod dose based on efficacy and safety for continued development

2.2. SECONDARY OBJECTIVES

The secondary objective is:

- To evaluate the effect of etrasimod on dysphagia symptoms in adult subjects with active EoE

CCI [redacted]

[redacted]

[redacted]

2.4. ESTIMANDS

The primary and secondary estimands to support regulatory decisions are described in Table 1 below. Supplementary analyses for the primary and secondary estimands will be performed in the Modified Full Analysis Set (mFAS), Completers, and Per Protocol Set, respectively.

List of Estimands

Estimand	Definition	Attributes			
		Population	Variable/ Endpoint	Intercurrent event handling strategy	Population-level summary measure
Primary Estimand	Efficacy of etrasimod on esophageal peak eosinophil count (PEC) at Week 16	Full Analysis Set	Percentage change from Baseline in esophageal PEC at Week 16	Subjects with missing data for any reason (including early discontinuation) or who initiated rescue medication and/or rescue procedure before the primary efficacy assessments prior to Week 16 will have subsequent PEC data set to missing from that point forward through Week 24. The missing data at Week 16 will be imputed using a multiple imputation (MI) method.	Differences between etrasimod (2 mg, 1 mg and any dose) and placebo in percent change from Baseline in esophageal PEC at Week 16

Key Secondary Estimand 1	Efficacy of etrasimod on Dysphagia Symptom Questionnaire (DSQ) at Week 16	Full Analysis Set	Change from Baseline in DSQ at Week 16	Same as the intercurrent event handling strategy for the primary estimand	Differences between etrasimod (2 mg, 1 mg and any dose) and placebo in change from Baseline in DSQ at Week 16
Key Secondary Estimand 2	Efficacy of etrasimod on esophageal PEC at Week 16	Full Analysis Set	Change from Baseline in esophageal PEC at Week 16	Same as the intercurrent event handling strategy for the primary estimand	Difference between etrasimod (2 mg, 1 mg and any dose) and placebo in change from Baseline in esophageal PEC at Week 16
Key Secondary Estimand 3	Efficacy of etrasimod on esophageal PEC at Week 16	Full Analysis Set	Proportion of subjects with esophageal PEC < 15 eos/hpf at Week 16	Subjects in any of the following cases will be treated as nonresponders: use of rescue medication and/or rescue procedure prior to Week 16; premature discontinuation for study due to any reason prior to Week 16; or with missing data at Week 16 for any reason. If subject had rescue medication and/or rescue procedure after Week 16, any assessment after the rescue medication and/or rescue procedure time will be set to missing.	Difference between etrasimod (2 mg, 1 mg and any dose) and placebo in the proportion of responders at Week 16

Key Secondary Estimand 4	Efficacy of etrasimod on esophageal PEC at Week 16	Full Analysis Set	Proportion of subjects with esophageal PEC ≤ 6 eos/hpf at Week 16	Same as intercurrent event handling strategy for key secondary estimand 3	Difference between etrasimod (2 mg, 1 mg and any dose) and placebo in the proportion of responders at Week 16

[a] Intercurrent events include: 1) initiate a rescue medication and/or rescue procedure for EoE, 2) have a rescue non-drug treatment before the efficacy assessment.
ES, endoscopic score; RB, rectal bleeding, SF, stool frequency; UC, ulcerative colitis.

3. STUDY DESIGN

3.1. GENERAL DESCRIPTION

This Phase 2 randomized, double-blind, multi-center study will evaluate the efficacy, safety, and PK of etrasimod compared with placebo for 24 weeks in adults with active EoE.

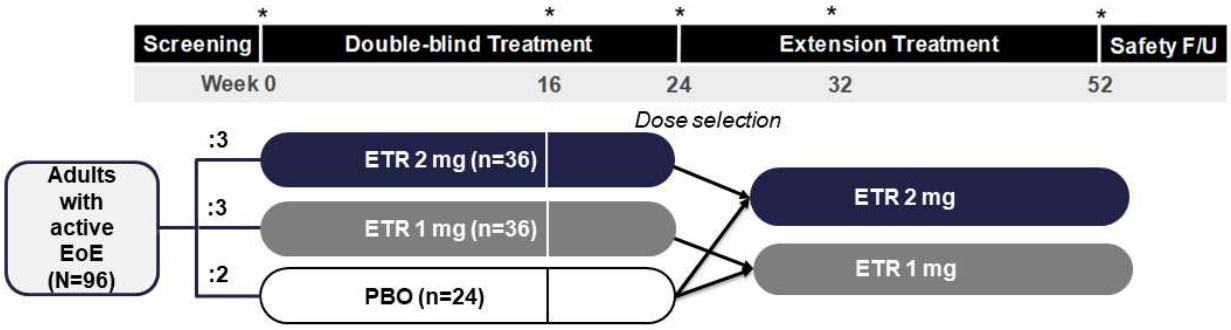
The study will consist of a Screening Period of up to 35 days, 24 weeks of double-blind treatment (Double-Blind Treatment Period), 28 weeks of active extended treatment (Extension Treatment Period), and 4 weeks of follow-up (Safety Follow-Up Period) for a total study duration of up to 61 weeks. Approximately 96 subjects were planned to be enrolled.

Eligible subjects will be randomized in a double-blinded fashion (3:3:2 ratio) to etrasimod 1 mg, etrasimod 2 mg, or matching placebo once daily. Randomization will be stratified by baseline history of esophageal dilation (yes/no) and concurrent proton pump inhibitor (PPI) therapy (yes/no).

All subjects who complete the Double-Blind Treatment Period and meet eligibility criteria for the Extension Treatment Period may enter the Extension Treatment Period. Subjects who were in the etrasimod 1 mg or etrasimod 2 mg groups in the Double-Blind Treatment Period will continue the same etrasimod dose in the Extension Treatment Period. Subjects who were in the placebo group during the Double-Blind Treatment Period will be re-randomized (1:1 ratio) to etrasimod 1 mg or etrasimod 2 mg at entry into the Extension Treatment Period.

Subjects will have follow-up visits at 2 and 4 weeks after the last dose of study treatment after Week 52 or the Early Termination (ET) visit.

Figure 1: Study Design



*Esophagogastroduodenoscopy
EoE, eosinophilic esophagitis; ETR, etrasimod; F/U, follow-up; PBO, placebo.
Numbers preceded by colons represent the randomization ratio.

3.2. SCHEDULE OF EVENTS

Schedule of events can be found in Appendix 1 of the protocol.

3.3. CHANGES TO ANALYSIS FROM PROTOCOL

- Not applicable.

4. PLANNED ANALYSES

The following analyses will be performed for this study:

- Main Analysis
- Final Analysis

4.1. MAIN ANALYSIS

A Main Analysis will take place for this study at the end of the Double-Blind Treatment Period when all subjects have completed Week 24 visit or discontinued from the study prior to Week 24. The results of which will be based on the unblinded treatment groups. All planned Main Study Analyses identified in this SAP will be performed by IQVIA Biostatistics following Pfizer Pharmaceutical Sponsor Authorization of this SAP, database unblinding, and Analysis Sets.

Derivations and definitions for the Main Analysis will be based on those required for the final analysis contained in this SAP, unless deviations are stated within the text. The list of outputs provided with the full set of output templates (planned for the final analysis) will highlight which of these outputs will also be provided for the Main Analysis.

After all subjects have completed the Double-Blind Treatment Period or discontinued from the study prior to Week 24, outstanding data queries have been resolved/closed, and the data have been cleaned and finalized, the sponsor will authorize breaking of the study blind for Double Blind Treatment Period and the Main Analysis of the data will be performed. No treatment assignment unblinded, or analyses completed until this SAP has been approved.

At the time of the main analysis, the sponsor staff will be unblinded to the initial treatment assignments at the beginning of Double-Blind Treatment Period, but participants, investigators and site staff will remain blinded. During the Extension Treatment Period, participants, investigators, and site staff will be aware that active treatment is being provided; however, they will remain blinded to the dose assignments. An unblinded team at IQVIA will receive the Double-Blind Treatment assignments and perform the Main Analysis

4.2. FINAL ANALYSIS

The Final Analysis will occur when all subjects have completed the Extension Treatment Period or discontinued from the study. All final, planned analyses identified in SAP will be performed by IQVIA Biostatistics following Pfizer Sponsor Authorization of this SAP, database unblinding, and Analysis Sets.

After all subjects have completed the study, outstanding data queries have been resolved/closed, and the data have been cleaned and finalized, the sponsor will authorize locking the database, unblinding for the second randomization schedule used for subjects receiving placebo in the Double Blind Treatment period and entering into the extension periodand the final analysis of the data will be performed. No database may be unblinded or analyses completed until this SAP has been approved.

Any, post-hoc, CCI [redacted] completed to support planned study analyses, which were not identified in this SAP, will be documented and reported in the clinical study report. Any results from these unplanned analyses will also be clearly identified in the text of the clinical study report. CCI [redacted]

Unless specified otherwise, the following treatment (treatment sequence) will be used for Double-blind Treatment Period and Extension Treatment Period, respectively:

List of treatments/treatment sequences

Study Period	Treatment / Treatment Sequence
Double-Blind Treatment Period	• etrasimod 1 mg • etrasimod 2mg • etrasimod any dose • Placebo
Extension Treatment Period	• etrasimod 1 mg – 1 mg • etrasimod 2mg – 2 mg • Placebo – etrasimod 1 mg • Placebo – etrasimod 2 mg

5. ANALYSIS SETS

Agreement and authorization of subjects included/excluded from each analysis set will be conducted prior to the unblinding of the study for the Main Analysis and Final Analysis.

5.1. PROCESS FOR ANALYSIS SET ASSIGNMENT

Before the database is unblinding for the Main Analysis and Final Analysis, the subjects excluded from each analysis set and the reason for exclusion will be sent to Pfizer for blinded review. The Clinical Protocol Deviation log file will be needed to define the Per Protocol Set for each analysis. This file will be incorporated into the analysis datasets and used, along with programmable deviations, to determine which subjects to include in the Per Protocol Set. Any changes to the analysis sets will occur before the database is unblinded. Pfizer will approve the final analysis sets which will be used for analyses prior to database unblinding. Once the analysis sets are approved, the subjects will be unblinded. Any additional analysis set assignments after database unblinding will also need to be approved.

5.2. SCREENED SET

The Screened Set will contain all subjects who provide informed consent for this study.

5.3. RANDOMIZED SET

The Randomized Set will contain all subjects who are randomized to study treatment.

5.4. FULL ANALYSIS SET

The Full Analysis Set (FAS) will consist of all randomized subjects in the Double-Blind Treatment Period who receive at least 1 dose of study treatment. Subjects will be summarized by treatment to which they were randomized, regardless of treatment actually received.

5.5. PER PROTOCOL SET

The Per Protocol Set will consist of all subjects in the FAS without major protocol violations on or before Week 16 that might affect the evaluation of the effect of study treatment on the primary endpoint. This set will be used in sensitivity analyses of the primary and secondary endpoints to evaluate the influences of major protocol violators and deviators on the primary results. Subjects will be excluded from the Per Protocol Set if they violate the eligibility criteria or significantly deviate from the study plan, including:

- Study treatment non-compliance (< 80 or > 120%)
- Received incorrect study treatment for > 7 days in total
- Used rescue medication or undergo rescue medical procedure that may affect primary efficacy endpoint at Week 16
- Missed Week 16 esophageal PEC in pre-specified analysis window per Section 6.4 for reasons other than discontinuing due to lack of efficacy

All protocol deviations will be reviewed in addition to the programmable deviations listed above (e.g., study treatment non-compliance) to determine if any deviation is significant enough to exclude a subject from the Per Protocol Set. All exclusion flags will be finalized before study unblinding. Subjects excluded from the Per Protocol Set will be listed along with the reason category above. Subjects will be summarized by treatment to which they were randomized, regardless of treatment actually received.

5.6. MODIFIED FULL ANALYSIS SET

The modified FAS (mFAS) will consist of all randomized subjects who receive at least 1 dose of study treatment in the Double-Blind Treatment Period and have a Baseline and at least 1 post-randomization measurement, In addition, subjects with baseline DSQ total score <8 will be excluded from mFAS for DSQ based endpoint analyses. Subjects will be summarized by treatment to which they were randomized, regardless of treatment actually received. Note that the mFAS can vary with endpoints since some subjects may have the needed data for inclusion in the mFAS for some endpoints, but not for other endpoints.

5.7. SAFETY SET

The Safety Set will be derived for both the Double-Blind Treatment Period and the Extension Treatment Period, respectively and includes all randomized subjects who receive at least 1 dose of study treatment during the specified treatment period. In Double-Blind Treatment Period, subjects will be analyzed according to treatment received In Extension Treatment Period, subjects will be analyzed according to the treatment sequence received, regardless of randomization. The Safety Set will be used for all safety analyses.

5.8. PHARMACOKINETIC SET

The Pharmacokinetic Set will include all subjects in the Safety Set with at least 1 quantifiable postdose PK measurement of etrasimod which is not impacted by protocol violations or events with potential to affect the PK concentration.

5.9. COMPLETERS SETS

The Completers Set (Week 16) will consist of all subjects in the FAS who complete Week 16 visit during the Double-Blind Treatment Period. Subjects will be summarized by treatment to which they were randomized.
The Completers Set (Week 24) will consist of all subjects in the FAS who complete Week 24 visit during the Double-Blind Treatment Period. Subjects will be summarized by treatment to which they were randomized.

5.10. EFFICACY EVALUABLE SET

The Efficacy Evaluable Set will be derived for the Extension Treatment Period only and includes all randomized subjects who receive at least 1 dose of study treatment and have at least one efficacy measurement in the Extension Treatment Period. Subjects will be analyzed according to treatment received in the Double-Blind Treatment Period and Extension Treatment Period (see treatments/treatment sequences in **Section 4.2** for more details).
Note that the Efficacy Evaluable Set can vary with endpoints since some subjects may have the needed data for inclusion in the Efficacy Evaluable Set for some endpoints, but not for other endpoints.

6. GENERAL CONSIDERATIONS

6.1. REFERENCE START DATE AND STUDY DAY

Study Day will be calculated from the reference start date and will be used to show start/stop day of assessments and events. Unless specified otherwise, reference start date is defined as the date of first dose (Day 1) and will appear in every listing where an assessment date or event date appears.

If the date of the event is on or after the reference date then:

Study Day = (date of event – reference date) + 1

If the date of the event is prior to the reference date then:

Study Day = (date of event – reference date)

In the situation where the event date is partial or missing, Study Day, and any corresponding durations will appear missing in the listings.

6.2. BASELINE

Unless otherwise specified, Baseline for the Double-Blind Treatment Period is defined as the last non-missing measurement taken prior to the first dose (including unscheduled assessments), and Baseline for the Extension Treatment Period is defined as the last non-missing measurement taken prior to the date of first dose in Extension Treatment Period at Week 24. If measurements include time (except for Quality of Life assessments (QoL)), the date/time will be used to define Baseline. Otherwise, only dates will be compared. QoL baseline is measured on Day 1 prior to the first dose. In the case where the last non-missing measurement and the date of first dose coincide and time is not collected, if the measurement was planned in the protocol to be done prior to the date of first dose, that measurement will be considered in defining Baseline.

6.3. RETESTS, UNSCHEDULED VISITS, AND EARLY TERMINATION DATA

For by-visit analyses and summaries, efficacy, safety, health-related quality of life, and CCI [redacted] (including scheduled, retests, unscheduled, and early termination) will be assigned to visits after the application of the windowing conventions described in Section 6.4. All measurements will be considered in summaries of abnormalities or worst-case values post-Baseline.

The visit windowing will be applied before missing data are imputed.

Listings will include scheduled, unscheduled, retest, and early discontinuation data.

6.4. WINDOWING CONVENTIONS

All scheduled study visits are defined relative to Study Day 1, the date of first dose. Scheduled visit windows are defined in Appendix 1 of the protocol. A windowing convention will be used to determine the analysis visit value for a given measurement and will be applicable for all by-visit summaries and analyses for efficacy and safety data. See Table 3 below for specific visit windows.

Visit Windows for Efficacy and Safety Analyses

For all efficacy (except for DSQ total score), CCI [redacted] health-related quality of life, safety labs and vital signs.	
Scheduled Study Visit (Protocol Scheduled Day)	Analysis Visit Window (Study Day)
Baseline (Double-Blind Treatment Period)	≤ 1, in addition to Study Day ≤ 1 , a measurement also needs to fulfill the baseline criteria identified in Section 6.2 in order to be considered as baseline measurement
Week 0 (Day 1)	1
Week 2 (Day 15 ± 3)	2 to 22
Week 4 (Day 29 ± 3)	23 to 43
Week 8 (Day 57 ± 5)	44 to 71
Week 12 (Day 85 ± 5)	72 to 99
Week 16 (Day 113 ± 5)	100 to 127
Week 20 (Day 141 ± 5)	128 to 155
Week 24 (Day 169 ± 5)	156 to 183
Week 28 (Day 197 ± 7)	184 to 211
Week 32 (Day 225 ± 7)	212 to 274
Week 46 (Day 323 ± 7)	275 to 344
Week 52 (Day 365 ± 7)	> 344

DSQ total score will be calculated every 2 weeks (week 0, 2, 4, 6, 8, 10, etc.) based on the scheduled study day starting on Day 1 For example, DSQ total score for Week 2, 4 and 6 will be calculated on Day 15, 29 and 43, respectively. .

For ECGs, OCT, and PFT.

Scheduled Study Visit (Protocol Scheduled Day)	Analysis Visit Window (Study Day)
Baseline (Double-Blind Treatment Period)	≤1
Week 0 (Day 1)	1
Week 24 (Day 169 ± 5)	71 to 267
Week 52 (Day 365 ± 7)	> 267

Windowing will be applied prior to any missing data calculations. The measurement that meets the Baseline criteria described in Section 6.2 will be labeled as "Baseline". Unscheduled or Screening visits that occurred before the Baseline visit will not be assigned an analysis visit. The early discontinuation visit will be eligible for allocation to an analysis visit. The 2- and 4-Week Follow-Up visits will not be included in the visit windows and will be summarized separately without any window applied.

If one or more results for a variable are assigned to the same analysis visit, the result with the date closest to the protocol scheduled day will be used in the analysis. If two measurements in the same analysis visit window are equidistant from the protocol scheduled study day, the earlier measurement will be used in the analysis. If multiple assessments are available on the same day, then the average of the assessment will be used in the analysis, except for laboratory and ECG data where the assessment at the earliest time of the same day will be used. If both central and local assessments of the same lab test are available on the same day, the central result will take precedence over the local result. If both central and local assessments of the same ECG test are available on the same day, the assessment of clinical significance by local reader will take precedence over the determination by a central reader.

6.5. STATISTICAL TESTS

The default significance level will be 0.05, confidence intervals will be 95%, and all tests will be 2-sided, unless otherwise specified in the description of the analyses.

6.6. COMMON CALCULATIONS

For quantitative measurements, change from baseline and proportions will be calculated as:

- Change from Baseline = Test Value at Visit X – Baseline Value

- Percent Change from Baseline = ((Test Value at Visit X – Baseline Value) / Baseline Value) * 100
- Proportion at Visit X = Number of subjects satisfying criteria at Visit X / Total number of subjects at Visit X

6.7. GENERAL STUDY INFORMATION

A general table with summary of study information will be generated, including date of first subject signed informed consent form (ICF), last subject visit date and database lock date. All analyses will be conducted using SAS® (version 9.4, SAS Institute Inc., Cary, NC).

CCI [redacted]

[redacted]

8. STATISTICAL CONSIDERATIONS

8.1. ADJUSTMENTS FOR COVARIATES AND FACTORS TO BE INCLUDED IN ANALYSES

The following covariates and factors, and/or endpoint Baseline measure are used in the analyses. For details of their inclusion in the models, see the specific analysis section.

- Baseline history of esophageal dilation (Yes/No)
- Concurrent PPI therapy (Yes/No)

If a subject was assigned the wrong stratum at randomization, their reported stratum based on the randomization system will be used for the statistical analyses. Analyses may be repeated using actual stratum based on eCRF as sensitivity analysis if randomization based on mis-stratification occurs in more than 10% of randomized subjects.

8.2. MULTICENTER STUDIES

- This study will be conducted by multiple investigators at multiple centers internationally.

Data from all sites will be pooled and statistical analyses will not be adjusted for investigational site, country, or geographic region.

8.3. MISSING DATA

Missing adverse event (AE) relationship to study drug and AE seriousness will be imputed as described in Section 17.1.1.2 and 17.1.3, respectively. Partial or missing AE start dates, Concomitant Medication (CM) start dates, and hospitalization dates will also be imputed as described in Appendix 2. No other missing safety data will be imputed.

Missing efficacy data will be handled as described in Sections 16.1.2, 16.2.2, and 16.3.2. In the primary analysis of the primary endpoint, PEC after rescue medication and/or rescue procedure uses will be set to missing; all missing data will then be imputed in the FAS using multiple imputation. In the main analyses of all binary responder-type endpoints, all subjects with missing data, regardless of reason for missingness, will be considered as nonresponders. In the main analysis of continuous or score endpoints, subjects with missing data will be handled using observed cases only, multiple imputation, or a linear mixed effect model.

8.4. MULTIPLE COMPARISONS/ MULTIPLICITY

For Double Blinded Treatment Period, hypotheses tests will be performed for efficacy endpoints for the comparison of etrasimod dose versus the placebo group. However, these hypothesis tests will not be controlled for type-I error rate. Un-adjusted p-values will be reported. The associated 95% confidence intervals will be reported.

For the Extension Treatment Period, no hypothesis tests between treatment groups will be performed; only descriptive statistics will be provided.

8.5. SUBGROUP ANALYSES

The following subgroups will be assessed for the primary, secondary, and CCI endpoints:

- Sex (Female, Male)
- Race (white, non-white)
- Age (≤ Median, > Median)
- History of Dilation (Yes, No)
- Duration of Disease (≤ Median, > Median)
- Concurrent PPI Therapy (Yes, No)
- Concurrent Atopic Comorbidities (Yes, No)

- Concurrent Elimination Diet (Yes, No)
- Baseline Esophageal PEC (≤ Median, > Median)
- Baseline DSQ Score (≤ Median, > Median)
- Baseline Eosinophilic Esophagitis Endoscopic Reference Score (EREFS) (≤ Median, > Median)
- Baseline Swallowed Topical Steroid Treatment Refractory Status (Yes, No)

Additional subgroups may be assessed, if deemed necessary. The medians will be derived based on the FAS for all subgroups cut at the median. It should be noted that the study was not designed to detect treatment differences with high statistical power within subgroups. If any subgroup includes < 5% of all subjects, no inferential statistics will be generated.

The actual stratum at Baseline will be used for all subgroup analyses.

9. OUTPUT PRESENTATIONS

0 shows conventions for presentation of data in outputs.

10. DISPOSITION AND PROTOCOL DEVIATIONS

All subjects who provide informed consent will be accounted for in this study. Inclusion criteria not met and exclusion criteria met will be listed.

Among the randomized subjects, the number and percent of subjects who completed/discontinued treatment, reasons off treatment, the number and percent of subjects who completed/discontinued the study, and reasons off study will be summarized in all randomized subjects. This summary will also be provided by region and country. The number and percent of subjects in each analysis set will be summarized in all randomized subjects. An additional summary of the number of subjects screened, the number of screen failed subjects, and reason for screen failure will be presented for all screened subjects. A listing of subjects whose blind was broken will be provided. The number of subjects whose visit was impacted by the COVID-19 pandemic per CRF will also be summarized by nominal visit.

During site monitoring, protocol deviations will be graded as Critical, Major, or Minor. According to ICH E3 and ICH E3(R1), important protocol deviations are a subset of protocol deviations that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a subject's rights, safety, or well-being (ICH 1995, ICH 2012). For example, important protocol deviations might include enrolling subjects in violation of key eligibility criteria designed to ensure a specific subject population or failing to collect data necessary to interpret primary endpoints, as this may compromise the scientific value of the study.

During the review of all reported deviations, important deviations related to study inclusion or exclusion criteria, conduct of the study, patient management, or patient assessment will be identified. Where relevant the importance of a potentially important protocol deviations will be assessed in the context of the study's estimands to evaluate potential impact. All important protocol deviations will be summarized for the FAS in the following categories in descending frequency by study treatment group.

All protocol deviations will be listed, including whether a deviation was impacted by the COVID-19 pandemic (Yes or No). Protocol deviation categories include but are not limited to the following:

- Informed Consent and Process
- Inclusion Criteria
- Exclusion Criteria
- Concomitant Medication Criteria
- Laboratory Assessment Criteria
- Study Procedures Criteria
- Blinding
- Patient Report Outcomes
- Safety
- Visit Schedule criteria
- Investigational Product Condition
- Subject Investigational Product Compliance
- Investigation Product (IP) Administration
- Efficacy Criteria
- Source Document Criteria
- Subject Discontinuation

Note, all protocol deviation categories above are based on the final PDMP that is completed after May 2022 except for Source Document Criteria, which is based on old PDMP prior to May 2022.

Important protocol deviations will also be summarized by whether they were impacted by COVID-19 pandemic (Yes or No). Additionally, a summary of missing endoscopy regardless of reason and missing endoscopy due to visit impacted by the COVID-19 pandemic will be presented by visit. The summary of important protocol deviations will be provided for Double-blinded Treatment period and Extension Treatment period, respectively.

The protocol deviation categories above are based on the latest version of PDMP and associated protocol deviation

categories prior to the unblinding of Double-blind Treatment period. If there is any update for these 2 documents regarding the protocol deviation categories after the unblinding of Double-blind Treatment period, the summary and listings regarding protocol deviation could be subject to be updated if deemed needed.

11. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

The following demographic data and Baseline characteristics will be summarized for the FAS:

- Age on consent (years)
- Sex
- Race
- Ethnicity
- Woman of childbearing potential (Yes or No)
- Height (cm)
- Weight (kg)
- Body mass index (BMI) (kg/m^2)
- Alcohol consumption (Yes or No)
- Tobacco use (Yes or No)

- This summary will be repeated for age, sex, race, ethnicity, or woman of childbearing potential for subjects who fail screening.

The following Baseline characteristics related to EoE will be summarized for the FAS:

- Primary EoE Symptom experienced (Dysphagia, Food Impaction, Specific food avoidance, other)
- Duration of EoE (years)
- History of Esophageal Dilation (Yes or No)
    - Subjects with more than 1 esophageal dilation (Yes or No)
    - Duration of Time Between the Last 2 Esophageal dilation procedures (months)
    - Impact of Dilation on EoE symptoms
    - Complications Related to Dilation Procedure (Yes or No)
- EoE Histopathologic Phenotype
- Biopsy of the esophagus previously assessed for PEC (Yes or No)
- Esophageal Level assessed for PEC: Proximal, Mild , Distal
- Endoscopic Findings
- Comorbid Conditions
- Type of Dietary Therapy
- Food Groups Eliminated from Food Elimination Diet
- Reason for Discontinuation of Diet

- Family History of EoE (Yes or No)
    - Relation
    - Type
- Baseline history of esophageal dilation (Yes or No) – Reported (used for stratification at randomization)
- Baseline history of esophageal dilation (Yes or No) – Actual (reported on the eCRF)
- Baseline history of esophageal dilation – Difference
- Concurrent PPI therapy (Yes or No) – Reported (used for stratification at randomization)
- Concurrent PPI therapy (Yes or No) – Actual (reported on the eCRF). The reference start date used to identify the Concurrent PPI therapy status is the date of randomization
- Concurrent PPI therapy – Difference
- Baseline PEC (eos/hpf)

- Prior EoE Treatment (Yes or No)
- Reason for Prior EoE Treatment Discontinuation

- Prior treatment for EoE will be summarized, including category of treatment, reason for discontinuation, and estimated duration (weeks) of treatment over the last 12 months.

12. DERIVATIONS

- Duration of EoE (years) = (Informed consent date – Date of diagnosis + 1) / 365.25
- Weight (kg) = Weight (lb) × 0.4536
- Height (cm) = Height (in) × 2.54
- Height (m) = Height (in) × 0.0254 = Height (cm) × 0.01
- BMI (kg/m^2) = Weight (kg) / Height $(m)^2$

13. MEDICAL HISTORY

Medical history will be collected on the medical history eCRF and coded using Medical Dictionary for Regulatory Activities (MedDRA, version 25.1 or higher). The version used to code medical history will be displayed in the outputs. All medical history will be summarized for the Safety Set by system organ class (SOC) and preferred term (PT).

14. MEDICATIONS

Medications will be captured on the Concomitant Medications eCRF and coded using the WHO Drug dictionary (WHODD01SEP2021 or later version). Refer to APPENDIX 2 for handling of partial dates for medications. In the case where it is not possible to define a medication as prior or concomitant, the medication will be classified by the worst case (i.e. concomitant).

- 'Prior' medications are medications which started and stopped prior to the first dose of study treatment.
- 'Concomitant' medications for the Double-Blind Treatment Period are medications which started prior to, on or after the first dose of study treatment AND ended on or after the date of first dose of study treatment or were ongoing at the end of the study.
- 'Concomitant' medications for the Extension Period are medications which started prior to, on or after the first dose of study treatment in the Extension Period AND ended on or after the date of first dose of study treatment in the Extension Period or were ongoing.

Concomitant medications will be summarized by Anatomical-Therapeutic-Chemical (ATC) Level 2 and Preferred Drug name for the Safety Set. Concomitant medications used for ongoing EoE will be flagged as such on the Concomitant Medications eCRF and summarized by ATC Level 2 and Preferred Drug name for the Safety Set. Concomitant medication summaries will be prepared for the End of Double-Blind Treatment Period and End of Extension Period, respectively.

Prior therapies for EoE, which will be captured on EoE Prior Therapies eCRF, will also be summarized by ATC Level 2 and Preferred Drug name for the Safety Set.

15. STUDY MEDICATION EXPOSURE

Exposure to study treatment in weeks will be summarized for the Safety Set for the End of Double-Blind Period, Extension Period, and overall, respectively. For overall exposure to study treatment in the Double-Blind Treatment Period, the date of first and last study treatment administration will be taken from the Dosing Administration for Double-Blind Period eCRF. For the overall exposure to study treatment in the Extension Period, the date of first and last study treatment will be taken from the Dosing Administration for Extension Period eCRF. Exposure for the entire study may also be summarized, in which case, the earliest study treatment date from both periods will be used for the first date and the latest study treatment date in both periods will be used for the last date. Interruptions, compliance, and dose changes are not taken into account for duration of exposure. Total subject-years on study will also be summarized.

Dose interruptions will be recorded on the Dosing Administration for Double-Blind Period and Dosing Administration for Extension Period eCRFs. Overdoses will be recorded on the Overdose eCRF. The frequency and percentage of subjects who had at least one dose interruption, who had at least one dose interruption of >7 days, who

had at least one dose interruption of >14 days, who had one overdose, and who had >1 overdose will be summarized for the Safety Set.

15.1. DERIVATIONS

Duration of exposure (weeks) = (date of last study treatment administration – date of first study treatment administration + 1) / 7. For subjects with missing study treatment end dates collected on the Dosing Administration eCRFs, their date of last study treatment administration will be imputed by the last date of all dosing administration start/stop dates recorded.

16. STUDY MEDICATION COMPLIANCE

Total number of tablets expected, total number of tablets taken, total number of tablets missed, overall compliance to study treatment, frequency and percentage of subjects with overall compliance of <80% or >120% are based on Dosing Administration in the EDC. They will be summarized for the Safety Set for both the Double-Blind Period and Extension Period.

16.1. DERIVATIONS

Compliance to study treatment is based on the Drug Accountability eCRF and will be calculated as the number of tablets taken (total dispensed – total returned) divided by the prescribed number of tablets expected during the treatment period, expressed as a percentage, see calculations below.

The total number of tablets expected is defined as the number of tablets that a subject is expected to have taken between their first and last study treatment administration and is numerically identical to the subject's overall study treatment exposure, since the medication is to be taken once daily. On any site visit day, the medication is to be held and taken at the site, after all predose assessments have been completed. For example, if a subject took their last dose of study treatment on Day 60 and returned to the site on Day 78 to return the study treatment bottle, then the total number of tablets expected would be 60, not 77.

- Overall Compliance to study treatment will be calculated as follows:

$$\frac{([\text{Total N of tablets dispensed in the study}] - [\text{Total N of tablets returned in the study}])}{[\text{Date of last dose}] - [\text{Date of first dose}] + 1} \times 100$$

If a bottle is not returned, all dispensed tablets will be assumed to be taken. For each subject, if a high percentage (> 25%) of bottles were not returned by a subject, additional analyses may be done where bottles not returned are excluded from the overall compliance calculation for the subject. In such analysis, the date of last dose or the date of last bottle return, whichever is earlier, will be used as the "date of last dose" in the calculation above.

Both scheduled and unscheduled study treatment dispensations will be used in the compliance calculation. Overall

compliance calculations will be performed for the Double-Blind Treatment Period and the entire study and will be used in determining inclusion/exclusion of subjects in the Per Protocol Set. For subjects discontinued prior to Week 16, their overall compliance for the 16 Week Treatment Period and the entire study will be identical. For subjects who remained in the study beyond Week 16, the 16 Week total tablets expected will be defined as Date of Week 16 visit – Date of first dose, since their last dose in the 16 Week Treatment Period should have been taken on the day before Week 16 visit.

17. EFFICACY OUTCOMES

17.1. PRIMARY EFFICACY

17.1.1. PRIMARY EFFICACY VARIABLE & DERIVATION

The primary efficacy variable is percent change from Baseline in esophageal PEC at Week 16. Esophageal PEC will be based on the central esophagogastroduodenoscopy (EGD) results. A decrease in esophageal PEC indicates improvement in disease.

17.1.2. MISSING DATA METHODS FOR PRIMARY EFFICACY VARIABLE

Subjects who 1) discontinue the study for worsening of disease, lack of efficacy, or adverse event related to EoE, 2) use a rescue medication for EoE, or 3) have rescue medical procedure during the study as confirmed after blinded review by clinical and medical team members before study unblinding will be considered to have a missing outcome in the analysis of all efficacy endpoints at any subsequent timepoints, including the primary endpoint. The rescue medications/rescue medical procedure will be identified by the Pfizer clinical team during blinded data review. The detailed information regarding the identification of rescue therapy is provided in Appendix 4. In scenarios 2 and 3 above, if a subject has an efficacy measurement collected after the initiation of rescue medication/rescue medical procedure, the observed data will be censored at the time of initiation and they will be considered as having a missing outcome. Subjects in all 3 scenarios above may still be included in the respective Per Protocol Set, provided they do not violate other criteria for Per Protocol Set. For example, they will be excluded from the respective Per Protocol Set if they initiate a prohibited medication before the Week 16 efficacy assessment that can affect efficacy of the study treatment and the indication is unrelated to EoE.

Analysis visits will be mapped as per Section 6.4 before any missing data imputation method is applied.

Subjects who initiate rescue medication will have PEC results after the start of rescue medication/rescue medical procedure set to missing. All missing PEC at Week 16 and Week 24 will then be imputed using multiple imputation under missing at random (MAR) procedure with a fully conditional specification (FCS) and predictive mean

matching for PEC as the primary missing data method. There are two different types of missing data patterns: non-monotone and monotone. A subject with missing data at any post-Baseline visit, and a non-missing data point after that visit is said to have intermittent missing data (non-monotone). A subject with missing data at any post-Baseline visit and at all subsequent visits is said to have monotone missing data.

17.1.3. MULTIPLE IMPUTATION UNDER MAR

Any missing PEC at the planned assessments will be imputed using multiple imputation under MAR. MAR assumes the missing value is independent of unobserved outcomes given observed data (i.e., subjects with missing PEC can be modeled based on subjects with observed PEC values) (Rubin 1987).

The following steps will be implemented:

Step 1:
Regardless of the arbitrary missing data pattern (i.e. monotone or non-monotone), a fully conditional specification (FCS) method with predictive mean matching for continuous variables will be used to impute the missing PEC data. The FCS method allows for separate conditional distributions for each imputed variable. The predictive mean matching approach creates a regression model using parameters sampled from the posterior distribution and then a predicted value for each missing value is computed. The missing value is replaced by randomly selecting an observation from a set of 'k' values that are the closest predicted values to the missing predicted value. Missing data imputation will be performed using the SAS PROC MI procedure. The number of imputations will be 40. A separate imputation model will be used for each treatment group. The stratification variables will be included in the imputation models.

Step 2:
Percent change from Baseline in PEC will be re-computed at Week 16 for each of the 40 imputed datasets.

Step 3:
The analysis of covariance (ANCOVA) model based on rank scores including treatment group, reported randomization stratification factors of Baseline history of esophageal dilation (Yes/No) and concurrent PPI therapy (Yes/No) as factors and Baseline eosinophil PEC as a covariate will be run for each of the 40 datasets to obtain 40 estimators of interest.

Step 4:
Use SAS PROC MIANALYZE to produce an overall pooled estimate (mean of 40 estimates) with its associated standard error, confidence interval, and pooled p-value.

There could be certain adjustments to the multiple imputation model due to unexpected data issues after unblinding

treatment. All post-unblinding modifications to the multiple imputation model or approaches to address missing data will be described in the Clinical Study Report (CSR).

Additionally, missing data will be imputed using the following missing data methods: Tipping point analysis and observed data only. More details are provided in Section 16.1.5.

17.1.4. PRIMARY ANALYSIS OF PRIMARY EFFICACY VARIABLE(S)

The following hypotheses (H01 and H02) for the primary endpoint will be tested:

H_{01}: The percentage change from Baseline in esophageal PEC at Week 16 is the same between etrasimod 2mg and placebo.
H_{11}: The percentage change from Baseline in esophageal PEC at Week 16 is different between etrasimod 2mg and placebo.
H_{02}: The percentage change from Baseline in esophageal PEC at Week 16 is the same between etrasimod 1mg and placebo.
H_{12}: The percentage change from Baseline in esophageal PEC at Week 16 is different between etrasimod 1mg and placebo.
All hypotheses will be tested at the 2 sided α level, 0.05.
The primary efficacy analysis will be performed for the FAS. Percent change from baseline in esophageal PEC will be summarized using the number of observations, mean, standard deviation (SD), median, minimum, and maximum. An analysis of covariance (ANCOVA) model based on rank scores due to the skewed distribution of the data and possible outliers will be used. The ANCOVA model will include treatment group and randomization stratification factors of Baseline history of esophageal dilation (Yes/No) and concurrent PPI therapy (Yes/No) as factors and Baseline eosinophil PEC as a covariate. Multiple results of the ANCOVA model will be expressed as least square (LS) means rank scores, LS mean rank scores difference from placebo, 95% confidence interval (CI) and p-value. For the treatment comparisons of etrasimod 1 mg vs. placebo and etrasimod 2 mg vs placebo, the treatment group factor in the ANCOVA model will include placebo, etrasimod 1 mg and etrasimod 2 mg. For the comparison of etrasimod any dose vs placebo, a separate ANCOVA model based on rank scores will be used, where treatment group factor will only include placebo and etrasimod any dose treatment. For the Extension Treatment Period, no pair-wise comparison will be performed; only descriptive statistics will be provided. In addition, descriptive summary of percentage change in esophageal PEC from baseline as well as difference between treatment groups will be presented by treatment group.

17.1.5. SUPPLEMENTARY ANALYSIS OF PRIMARY EFFICACY VARIABLE(S)

The primary model as described in Section 16.1.3 will be repeated using the Completers Set and Per Protocol Set as

supplementary analyses. Subgroup analyses will also be performed for the primary model.

17.1.6. SENSITIVITY ANALYSIS OF PRIMARY EFFICACY VARIABLE(S)

Sensitivity analyses will be implemented to explore different types of missing data approaches.

17.1.6.1. Tipping Point Analysis

A tipping point approach based on multiple imputation of values through Week 16 will be used with a specified adjustment (referred to as delta adjustment or shift) applied to values imputed under a MNAR-based imputation model. To find a tipping point, a series of imputations will be performed with increasing values of delta.
The goal is to evaluate the plausibility of the assumed expected values for missing outcomes on each treatment group under which the conclusions change, ie, under which values of delta there are no longer evidence of significant treatment difference.
The following steps will be implemented:
Step 1:
Same as Step 1 described in 16.1.2 above.

Step 2:
Apply a shift (i.e. additive delta adjustment) to the imputed PEC at each visit in both treatment groups. Percent change from Baseline in PEC will be re-computed at Week 16 for each of the 40 imputed datasets.

Step 3:
Same as Step 3 described in 16.1.2 above.

Step 4:
Same as Step 4 described in 16.1.2 above.

Note, imputations with delta adjustment will be performed with progressively increasing delta values until a tipping point is reached based on the significance of the ANCOVA test based on rank scores. A tipping point will correspond to the smallest value of delta for which the primary hypothesis is no longer rejected. The shift parameter (delta) for the active treatment group will take increasing values from 0 with increments of 40 (maximum PEC of X), representing adjustments in towards worse outcomes (i.e., after dropout, on average, subjects on treatment will have PEC worsen). For placebo arm, delta adjustments will decrease with increments of 40, representing adjustments in the direction of better outcomes. The adjustments in each treatment group will be continued until the ANCOVA test based on rank scores for the primary hypotheses are no longer significant at the 2sided α level, 0.05. The corresponding shift parameters in both treatment groups at the tipping point will be reported along with the LS means rank scores, LS mean rank scores difference from placebo, 95% CI, and p-values from each ANCOVA test based on rank scores.

Results will be plotted as a heat map on a two-dimensional plot with axes corresponding to the delta values used in each treatment group respectively, and different colors will be used to represent the magnitude of the 2-sided p-values corresponding to analysis with each combination of delta values. The clinical interpretation about the plausibility of the assumptions underlying the tipping point will be provided in the CSR.

17.1.6.2. Observed Data Only

This analysis will be performed using the mFAS. No imputation for missing data will be performed.

17.2. SECONDARY EFFICACY

The secondary efficacy analyses will be performed for the FAS.

17.2.1. SECONDARY EFFICACY VARIABLES & DERIVATIONS

The secondary efficacy analyses will be performed for the FAS. The secondary efficacy endpoints are:

- Change from Baseline in DSQ total score at Week 16
- Change from Baseline in Esophageal PEC at Week 16.
- Proportion of Subjects with Esophageal PEC < 15 eos/hpf at Week 16
- Proportion of Subjects with Esophageal PEC ≤ 6 eos/hpf at Week 16

Change from Baseline in DSQ total score at Week 16 will be based on daily DSQ score from subject daily Diary data. The DSQ total score and daily DSQ score calculations are provided below. An increase in DSQ total score indicates worsening of dysphagia symptoms.

Calculation Approach	DSQ Total Score	Daily DSQ Score
14-day period approach[a]	DSQ total Score = (Sum of daily DSQ score with 14-day period preceding the scheduled visit day) ×14 / (Number of days with diaries reported with nonmissing data within 14-day period). Shifting rule when calculating DSQ total score at each visit based on the number of reported diary entries within the 14-day period prior to the visit: Baseline visit : If there are less than 8 reported diaries that are available within the 14-day period prior to baseline, then the most recent 8 reported diaries in a consecutive 14-day period would be used to calculate the DSQ total core. In order to determine the most recent 14-day period with the minimum number of reported diaries, it may be necessary to make adjustments in some cases by shifting to earlier diary entries. Such 14-day periods cannot be shifted for more than 7 days (not including the study visit day). After shifting 7 days, if there are still less than 8 reported diaries	Daily DSQ Score= sum of answers to Questions 2 and 3 in the daily DSQ questionnaire Daily DSQ Score is set to missing if the diary is not reported or if the diary is reported but the answer to Question 1 is “No” in daily DSQ questionnaire. In this case the diary is considered as missing Daily DSQ Score is set to zero if the answer to Question 2 is “No” in daily DSQ questionnaire

	within the 14-day period, then the DSQ total score for baseline visit will be set to missing. Week 2 visit: If there are less than 8 reported post-baseline diaries are available within the 14-day period prior to week 2 visit, then DSQ total score will be set to missing without any shift rule applied. All other post baseline visits after Week 2: If there are less than 8 reported diaries that are available within the 14-day period prior to baseline, then the most recent 8 reported diaries in a consecutive 14-day period would be used to calculate the DSQ total core. In order to determine the most recent 14-day period with the minimum number of reported diaries, it may be necessary to make adjustments in some cases by shifting to earlier diary entries. Such 14-day periods cannot be shifted for more than 14 days (not including the study visit day). After shifting 14 days, if there are still less than 8 reported diaries within the 14-day period, then the DSQ total score for baseline visit will be set to missing. The DSQ total score range is 0 to 84	
14-day period: worst case scenario approach[b]	DSQ total score calculation: same as 14-day period approach Shifting rule: same as as 14-day period approach The DSQ total score range is 0 to 84	Daily DSQ Score= sum of answers to Questions 2 and 3 in the daily DSQ questionnaire Missing Daily DSQ Score handling approach is the same as 14-day approach except that: Daily DSQ Score is set to (maximum value of) 6 if the answer to Question 1 is “No” in daily DSQ questionnaire and the answer for Food Avoidance Question (FAQ) is “Because of eosinophilic

		esophagitis symptoms" Daily DSQ Score is set to zero if the answer to Question 2 is "No" in daily DSQ questionnaire
7-day period approach[b]	DSQ total Score = (Sum of daily DSQ score with 7-day period preceding the scheduled visit day) ×7 / (Number of days with diaries reported with nonmissing* data within 7-day period). Shifting rule when calculating DSQ total score at each visit based on the number of reported diary entries within the 7-day period prior to the visit: Baseline visit and each post baseline visit: If there are less than 4 reported diary entries that are available within the 7-day period prior to the visit, then the most recent 4 reported diary entries in a consecutive 7-day period would be used to calculate the DSQ total core. In order to determine the most recent 7-day period with the minimum number of reported diaries, it may be necessary to make adjustments in some cases by shifting to earlier diary entries. Such 7-day periods cannot be shifted for more than 7 days (not including the study visit day). After shifting 7 days, if there are still less than 4 reported diary entries within the 7-day period, then the DSQ total score for baseline visit will be set to missing. The DSQ total score range is 0 to 42	Same as 14-day period approach

[a] Primary approach to calculate DSQ total score

[s] Sensitivity approach to calculate DSQ total score

Change from Baseline in esophageal PEC at Week 16 will based on the central EGD results.

Proportion of subjects with esophageal PEC < 15 eos/hpf and ≤ 6 eos/hpf at Week 16, will be based on central reading of the EGD results. Subjects who achieve esophageal PEC < 15 eos/hpf and ≤ 6 eos/hpf will be referred to as responders, respectively. Subjects who do not achieve esophageal PEC < 15 eos/hpf and ≤ 6 eos/hpf will be referred to as nonresponders, respectively.

Planned Analyses by Endpoint

Endpoint (at Week 16)	ANCOVA[a]/ Weighted MH	Sensitivity Analyses: MI	Sensitivity Analyses: Tipping Point Analysis	Completer Sets and Per Protocol Set	Linear mixed effects model
Key primary endpoints					
Percent Change from baseline in esophageal PEC	X[c]	X[e]	X[c]	X[c]	
Key secondary endpoints					
Change from baseline in DSQ total score		X[e, f]	X[e, f]	X[e, f]	X[b, f, g]
Change from baseline in esophageal PEC	X[c]	X[e]	X[c]	X[c]	
Proportion esophageal PEC < 15 eos/hpf	X[d]	X[e]		X[d]	
Proportion esophageal PEC <=6 eos/hpf	X[d]	X[e]		X[d]	

[a] No rank transformation will be applied

[b] No missing data methods will be implemented since missing data will be handled in the linear mixed effect model.

[c] Subjects who initiate rescue medication/rescue procedure will have assessments collected after the rescue medication/rescue procedure is initiated set to missing.

[d] Subjects who initiate rescue medication/rescue procedure will have assessments collected after the rescue medication/rescue procedure is initiated set to missing. Missing data for any reason will be considered as nonresponder. Adjusted difference and p-values are based on the estimated common risk difference using the MH weights

[e] Subjects who initiate rescue medication/rescue procedure will have assessments collected after the rescue medication is initiated set to missing. Missing data will be imputed using multiple imputation.

[f] primary approach (14-day period approach) will be used to calculate DSQ total score

[g] sensitivity approaches (14-day period: worst case scenario approach and 7-day period approach) will be used to calculate DSQ total score.

In addition, some of the additional analyses above will be conducted as follows:
Percent change from baseline in esophageal PEC at Week 16 above will be repeated at week 24 using the same ANCOVA model.
Change from baseline in DSQ total score at Week 16 above will be repeated for week 2, 4, 6, 8, 10, 12, 14, 20 and 24 in Double-Blind Treatment Period, using the same linear mixed effects model for change from baseline in DSQ total score at Week 16.

17.2.2. Missing Data Methods for Secondary Efficacy Variable(s)

The same missing data methods used for the primary endpoint described in Section 16.1.2 will be used for the continuous secondary efficacy endpoints. For proportion-based secondary endpoints, subjects with missing data for any reason or took rescue medication and/or rescue procedure will be considered a nonresponder.

17.2.3. Primary Analysis of Secondary Efficacy Variables

17.2.3.1. Change from Baseline in DSQ total score and Esophageal PEC at Week 16

Analysis of the secondary efficacy variables change from Baseline in DSQ total core at Week 16 will use a linear mixed effect model. The linear mixed effect model will include treatment group, visit, interaction of treatment-by-visit, Baseline history of esophageal dilation (Yes or No), and concurrent PPI therapy (Yes or No) as factors and Baseline DSQ total score as a covariate using an unstructured covariance matrix. If there is a convergence issue in fitting the unstructured covariance, a compound symmetry covariance structure will be used. The Kenward-Rogers method will be used to estimate the denominator degrees of freedom. Change from Baseline in Esophageal PEC at Week 16 will use an ANCOVA model including treatment group and randomization stratification factors as factors and Baseline eosinophil PEC as a covariate. Eosinophil PEC assessments after rescue medication and/or rescue procedure use will be set to missing. LS means at visit and LS mean differences between treatment groups (i.e. etrasimod 1 mg vs placebo, etrasimod 2 mg vs placebo, and etrasimod pooled vs placebo) with p-values and corresponding 95% CIs will be reported. For the individual etrasimod summaries, treatment group will include placebo, etrasimod 1 mg and etrasimod 2 mg. For the pooled etrasimod summary, treatment group will include placebo and pooled etrasimod.

17.2.3.2. Proportion of Subjects with Esophageal PEC < 15 and ≤ 6 eos/hpf at Week 16

Analysis of the secondary efficacy variables proportion subjects with esophageal PEC < 15 eos/hpf at Week 16 and proportion of subjects with esophageal PEC ≤ 6 eos/hpf at Week 16 will be performed for randomization stratification factors of baseline history of esophageal dilation (Yes or No) and concurrent PPI therapy (Yes or No). Subjects who use rescue medication or have a rescue dilation, or with missing data at Week 16 for any reason will be considered nonresponders. The stratified analyses will be performed for randomization stratification factors. Results will be expressed as the number and percentage of subjects who achieved the goal, risk difference in

proportion between treatment groups achieving the goal using Mantel-Haenszel weights, the associated 95% CI and p-value. If there are zero responders in both treatment groups in any stratum, the difference in proportions and associated inferences can still be made.

17.2.4. Supplementary Analysis of Secondary Efficacy Variables

The primary analysis of the secondary efficacy variables as described in Section 16.2.3 will be repeated using the Completers Set and Per Protocol Set as supplementary analyses. Subgroup analyses will also be performed.

17.2.5. Sensitivity Analysis of Secondary Efficacy Variables

The same sensitivity analyses for the primary endpoint described in 16.1.5 will be performed for the continuous secondary efficacy endpoints.

For proportion-based secondary endpoints, a sensitivity analysis using the primary endpoint multiple imputation method described in 16.1.2 will be performed for the FAS.

CCI [redacted]

[redacted]

[redacted]

CCI

- [redacted]
- [redacted]
- [redacted]
- [redacted]

CCI [redacted]

[redacted]

CCI [redacted]

CCI [redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				
[redacted]				

[redacted]

[redacted]

CCI [redacted]

[redacted]

CCI [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

CCI [redacted]

[redacted]

[redacted]

18. SAFETY OUTCOMES

All outputs for safety outcomes will be based on the Safety Set. There will be no statistical comparisons between the treatment groups for safety data, unless otherwise specified. Subjects will be analyzed according to treatment received in the Double-Blind Treatment Period and Extension Treatment Period. Unless otherwise specified, the safety analyses will be conducted for Double-Blind Treatment period and Extension Treatment period, respectively, when its applicable.

18.1. ADVERSE EVENTS

Adverse Events (AEs) will be coded using MedDRA (version 25.1 or higher). Treatment-emergent adverse events (TEAEs) are defined as follows for Double-Blind Treatment period and Extension Treatment period, respectively:

- Double-Blind Treatment period: TEAE is defined as AEs that started or worsened in severity on or after the first dose of study treatment (etrasimod 1 mg, 2 mg or placebo) in Double-Blind Treatment period.
- Extension Treatment period: TEAE is defined as AEs that started or worsened in severity on or after the first dose of study treatment (etrasimod 1 mg or 2 mg) in Extension Treatment period.

See APPENDIX **2** for handling of partial dates for AEs for the purpose of assigning treatment-emergent flags. In the case where it is not possible to define an AE as treatment-emergent or not, the AE will be classified as treatment-emergent.

An overall summary of number of subjects within each of the categories described in the sub-section below, will be provided as specified in the templates. Listings will include TEAEs and Non-TEAEs.

18.2. ALL TEAEs

All TEAEs will be summarized by SOC and PT. This summary table and all other TEAE summaries by SOC and PT will be presented by descending frequency in the etrasimod pooled (1 mg and 2 mg) group. All AEs, regardless of treatment emergent status, will be included in an AE listing. Additionally, a listing of other AE details as collected on the CRF will be presented.

18.2.1. SEVERITY

Severity is classified as Grade 1: Mild, Grade 2: Moderate, Grade 3: Severe, Grade 4: Life-Threatening, Grade 5: Death Related to AE, using the Common Terminology Criteria for Adverse Events, version 5.0 (CTCAE v5.0).

All TEAEs will be summarized by SOC, PT, and maximum severity, with SOC and PT presented by descending frequency. If a subject reports a TEAE more than once within a SOC/PT, the AE with the worst-case severity will be used in this summary.

18.2.2. Relationship to Study Treatment

Relationship is classified as "not related", "unlikely related", "probably related", or "related" by the Investigator.

All related TEAEs will be summarized by SOC and PT. A "related TEAE" for the purpose of this summaries is defined as a TEAE with relationship to study drug of "probably related" or "related".

All TEAEs will be summarized by SOC, PT, and highest relationship (as reported on the eCRF, not grouped), with SOC and PT presented by descending frequency. If a subject reports a TEAE more than once within a SOC/PT, the AE with the worst-case relationship to study treatment will be used in this summary. TEAEs with a missing relationship to study treatment will be regarded as "Related" to study treatment for summary tabulation purpose only.

18.2.3. TEAEs Leading to Discontinuation of Study Treatment

TEAEs leading to permanent discontinuation of study treatment will be identified by action taken being recorded as "Drug withdrawal" on the Adverse events eCRF.

All TEAEs leading to discontinuation of study treatment will be summarized by SOC and PT. A listing of all TEAEs leading to discontinuation of study treatment will also be presented.

18.2.4. Serious Adverse Events

Serious adverse events (SAEs) are those events recorded as "Serious" on the Adverse Event eCRF.

All serious TEAEs will be summarized by SOC and PT. If the seriousness is missing, the AE will be considered as "Serious". A serious TEAE listing will also be presented. All non-serious TEAEs will also be summarized by SOC and PT.

18.2.5. TEAEs Leading to Death

TEAEs leading to Death are those events which are recorded with an outcome as "Fatal" on the Adverse Events eCRF. All TEAEs leading to death will be listed.

18.2.6. TEAEs OF SPECIAL INTEREST

Categories of Targeted Medical Events (TMEs) and a list of preferred terms associated with these TME categories were developed based on the mechanism of action of etrasimod, prior experience with other agents acting via a similar mechanism, and disease-specific clinical judgment. In addition, where standard testing has been implemented to screen for potential AEs(eg, electrocardiograms, spirometry, serum transaminases, etc.), the relevant data will be reviewed to identify potential cases of AESI that investigators may not have identified and to provide quantitative data for AESIs. The proposed candidate terms will be reviewed to identify which events reflect AESI.

TEAEs of special interest will be summarized by category, subcategory, and PT by descending frequency by treatment group. Categories and subcategories of TEAEs of special interest are the following:

- Cardiovascular Events
  - Bradycardia
  - AV conduction delay
  - Hypertension
- Infections
  - Severe infections
  - Opportunistic infections (Narrow)
  - Herpes simplex and herpes zoster
- Macular Edema
- Pulmonary Disorders
  - Airflow obstruction (forced expiratory volume in 1 second [FEV_1], forced vital capacity [FVC])
  - Decrease gas exchange (diffusing capacity of the lungs for carbon monoxide [DLCO])
- Liver Injury
  - Liver transaminases elevation
  - Bilirubin elevation
  - Elevated BP
- Posterior reversible encephalopathy syndrome (PRES)
- Malignancies

For the TEAE of special interest, the Exposure-adjusted incidence rate (EAIR) will also be provided for the following group based on the subject in Safety Set in Double-Blinded Treatment Period

Group 1: Subjects who take etrasimod 2 mg anytime during the course of study. This includes the subjects who

- receive the etrasimod 2 mg in Double-Blind Treatment period and continue receiving Etrasimod 2 mg in Extension Treatment period, or
- receive the etrasimod 2 mg in Double-Blind Treatment period and discontinue early without entering Extension Treatment period, or
- receive the placebo in Double-Blind Treatment period and receive etrasimod 2 mg in Extension Treatment period

Group 2: Subjects who take etrasimod 1 mg anytime during the course of study. This includes the subjects who

- Subjects who take etrasimod 1 mg anytime during the course of study. This includes the subjects who
- receive the etrasimod 1 mg in Double-Blind Treatment period and continue receiving etrasimod 1 mg in Extension Treatment period, or
- receive the etrasimod 1 mg in Double-Blind Treatment period and discontinue early without entering Extension Treatment period, or
- receive the placebo in Double-Blind Treatment period and receive etrasimod 1 mg in Extension Treatment period

Group 3: Subjects who take placebo in Double-Blind Treatment period

The EAIR for a given TEAE of special interest is calculated as the number of subjects with the TEAE of special interest divided by the total exposure in subject-years. The 95% confidence interval (CI) of the EAIR will be provided along with EAIR. Some notes for the exposure and AE used in EAIR calculation are provided as follows

Group	Note for Exposure used in EAIR calculation	Note for TEAE of special interest used in EAIR calculation
1)	sum of either time (year) from first dose of etrasimod 2 mg to the onset of first such event for those who experienced the TEAE of interest, or time (year) from first dose of etrasimod 2 mg to last participation for those who did not experience the TEAE of special interest.	For a placebo subject who receive 2 mg in Extension Treatment period, only the AE that started or worsened in severity on or after the first dose of etrasimod 2 mg in Extension Treatment period will be included in the EAIR for the subject
2)	sum of either time (year) from first dose of etrasimod 1 mg to the onset of first such event for those who experienced the TEAE of interest, or time (year) from	For a placebo subject who receive 1 mg in Extension Treatment period, only the AE that started or worsened in severity on or

	first dose of etrasimod 1 mg to last participation for those who did not experience the TEAE of special interest.	after the first dose of etrasimod 1 mg in Extension Treatment period will be included in the EAIR for the subject
3)	sum of either time (year) from first dose of placebo to the onset of first such event for those who experienced the TEAE of interest during Double-Blind Treatment period, or time (year) from first dose of placebo to last participation for those who did not experience the TEAE of special interest. during Double-Blind Treatment period	For a placebo subject in Double Blind Treatment period, only the AE that started or worsened in severity on or after the first dose of placebo in Double Blind Treatment period will be included in the EAIR for the subject

18.2.7. Overall Summary of Adverse Events

In addition to the summaries above, an overview of TEAEs will be summarized (not broken down by SOC or PT) by number and frequency of subjects and by number of AEs:

- Any TEAEs
  - Any related TEAEs*
- Any serious TEAEs
  - Any related serious TEAEs*
- TEAEs leading to death
- TEAEs leading to study drug discontinuation
  - Related TEAEs leading to study drug discontinuation
- TEAEs leading to study drug interruption
  - Related TEAEs leading to study drug interruption*
- TEAEs by maximum severity
  - Related TEAEs by maximum severity*
- TEAEs by relationship to study drug

* “Related TEAEs” refers to TEAEs related or probably related to study drug or are missing relationship.

18.3. Deaths

Information collected about deaths (e.g., date of death, primary cause of death) will be presented in a data listing, as described in Section 18.2.5.

18.4. LABORATORY EVALUATIONS

No local laboratory assessments will be used in any summaries except for lipid panel and thyroid panel tests. No local laboratory assessments will be used to derive maximum/minimum/worst value. Local laboratory assessments will be listed.

18.4.1. SAFETY LABORATORY EVALUATIONS

Hematology, serum chemistry, and coagulation are analyzed and reported by central laboratory and sometimes by local laboratory. Results out of reference range are flagged by the performing laboratory (e.g., low, high). A full list of laboratory assessments to be included in the outputs is included in Table 3 of protocol amendment v3.0 (04Jan2023).

Presentations will use SI Units. Quantitative laboratory measurements reported as "< X" or "> X", where X may be the lower limit of quantification (LLQ) or the upper limit of quantification (ULQ), respectively, will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e. as "< X" or "> X" in the listings. For urinalysis, only pH and specific gravity are considered as quantitative tests. If local and central laboratory assessments are available from the same day, central laboratory assessments will be used in the summary as per Section 6.4.

The following summaries will be provided for laboratory data:

- Value and change from Baseline by visit (for hematology, TBNK panel, serum chemistry, quantitative urinalysis [pH and specific gravity], and coagulation)
- Incidence of abnormal values according to laboratory reference ranges by visit
- Incidence of lymphocytes <0.2 x 10^9/L, 0.5 x 10^9/L, or neutrophils < 0.5 x 10^9/L, 1 x 10^9/L at end of treatment and anytime in the study
- Change and Percent Change from Baseline in Absolute Lymphocyte Count (ALC) at Weeks 16, 24, 32, and 52
- Shift from Baseline to end of treatment according to laboratory reference range (for quantitative measurements and categorical measurements)
- eDISH plots for the following laboratory assessments:
- Maximum Alkaline Phosphatase vs. Same-Day Total Bilirubin
- Maximum Alkaline Phosphatase vs. Same-Day Total Bilirubin Adjusted for Elevated Baseline
- Maximum Alanine Aminotransferase vs. Same-Day Total Bilirubin
- Maximum Alanine Aminotransferase vs. Same-Day Total Bilirubin Adjusted for Elevated Baseline
- Maximum Aspartate Aminotransferase vs. Same-Day Total Bilirubin
- Maximum Aspartate Aminotransferase vs. Same-Day Total Bilirubin Adjusted for Elevated

Baseline

If both central and local assessments of total bilirubin are available on the same day, the central result will take precedence over the local result in the eDISH plot. If the maximum AST/ALT/GGT assessment occurs at two different dates, the assessment with the higher accompanying Bilirubin value will be used. Two versions of the eDISH plots will be presented, with one showing values as multiples of upper limit of normal, and the other one showing values as multiples of upper limit of normal, or subject's baseline, whichever is higher.
Only laboratory tests completed by 50 or more subjects in the SAF will be presented in the summaries. All laboratory tests will be listed.
Subject's laboratory assessments at all timepoints will be listed in chronological order. Values outside of the laboratory reference range will be flagged. Values obtained from local laboratory will be flagged. Listing of lymphocytes and neutrophils over time in subjects ever with lymphocytes $<0.5 \times 10^9$/L or neutrophils $< 1 \times 10^9$/L will also be provided.

18.4.2. PREGNANCY TESTS

Urine beta-human chorionic gonadotropin (β-hCG) and/or serum β-hCG pregnancy tests are performed in female subjects of childbearing potential. All pregnancy test results will be listed in chronological order.

18.4.3. OTHER SCREENING LABORATORY ASSESSMENTS

Laboratory assessments for virology, T, B, NK, CD4+ T cell and CD8+ T cell counts (TBNK) and Urine drug screening will be listed. Analyses of genetics and CCI [redacted] data based on samples collected in subjects who provided consent will be described in a separate plan.

18.5. ECG EVALUATIONS

ECGs are recorded on a 12-lead ECG machine and read locally and centrally. The following ECG parameters will be reported for this study:

HR (bpm)
PR Interval (ms)
RR Interval (ms)
QRS Interval (ms)
QT Interval (ms)
QTcF Interval (ms)
Overall interpretation of ECG (Investigator's judgment):

- Normal
- Abnormal, Not Clinically Significant (Abnormal NCS)

- Abnormal, Clinically Significant (Abnormal CS)

Overall interpretation of ECG (central reader):

- Normal
- Abnormal, Not Clinically Significant (Abnormal NCS)
- Abnormal, Clinically Significant (Abnormal CS)

The following summaries will be provided for ECG data:

- Value and change from Baseline by visit (for quantitative measurements)
- Incidence of markedly abnormal values (defined in Section 17.4.1) by visit
- Shift in normal/abnormal NCS/abnormal CS in the overall interpretation (by investigator) from Baseline to end of treatment and to the worst-case post-Baseline
- Shift in markedly abnormal categories from Baseline to post-Baseline by visit

All ECG results, including first dose cardiac monitoring, will be listed. Discharge criteria for first dose cardiac monitoring will also be listed.

A listing of all ECG assessments over time in subjects meeting markedly abnormal criteria will also be provided. For each subject, only ECG parameters ever meeting markedly abnormal criteria will be included.

18.5.1. ECG MARKEDLY ABNORMAL CRITERIA

Markedly abnormal quantitative ECG measurements will be identified in accordance with the following predefined markedly abnormal criteria:

- Absolute values in QT and QTcF:
  - ≥ 450 ms (male) or ≥ 470 msec (female) in QTcF
  - > 500 ms in QT
- Change from Baseline in QT and QTcF:
  - > 30 ms increase from Baseline
  - > 60 ms increase from Baseline

In shift tables, subjects will be classified according to the binary category for each parameter and the predefined markedly abnormal criterion (i.e., Markedly abnormal vs. Not markedly abnormal, with markedly abnormal defined in the footnote).

18.6. VITAL SIGNS

The following Vital Signs measurements will be reported for this study:

- Systolic Blood Pressure (mmHg)
- Diastolic Blood Pressure (mmHg)
- Heart Rate (bpm)
- Respiratory Rate (resp/min)
- Temperature (ºC)
- Weight (kg)
- Height (cm) [Screening only]

The following summaries will be provided for vital signs data:

- Value and change from Baseline by visit
- Value and change from predose on Day 1 and Week 24 (as reported on the eCRF) by timepoint
  - For heart rate only, also include value and change from predose to minimum postdose heart rate on Day 1 in the same table
- Incidence of markedly abnormal values (defined in section 17.5.2) by visit
- Listing of subjects meeting markedly abnormal criteria
- Incidence of minimum heart rate on Day 1 by postdose timepoint (1, 2, 3, 4, and > 4 hours postdose, and Day 1 overall) and heart rate interval (≥ 65, 60 to 64, 55 to 59, 50 to 54, 45 to 49, 40 to 44, < 40 bpm)
  - If minimum heart rate is attained at multiple postdose timepoints, only the earliest timepoint will be counted in this incidence summary
- Time to minimum heart rate on Day 1 by planned hourly timepoint (if minimum heart rate is attained at multiple postdose timepoints, only the earliest timepoint will be counted). Actual time elapsed from first dose to minimum heart rate on Day 1 will also be summarized using descriptive statistics (n, mean, SD, median, minimum, and maximum) in the same table.

All vital signs, including first dose cardiac monitoring, will be listed. Discharge criteria first dose cardiac monitoring will also be listed. A listing of all vital signs assessments over time in subjects meeting markedly abnormal criteria will also be provided. For each subject, only vital sign parameters ever meeting markedly abnormal criteria will be included.

For subjects with extended monitoring, systolic blood pressure, diastolic blood pressure and heart rate values and change from predose on Day 1 will be listed.

18.6.1. VITAL SIGNS SPECIFIC DERIVATIONS

- Temperature (ºC) = (5/9) (Temperature (ºF) – 32)

18.6.2. VITAL SIGNS MARKEDLY ABNORMAL CRITERIA

Markedly abnormal quantitative Vital Signs measurements will be identified in accordance with the following predefined markedly abnormal criteria:

Table 5: Markedly Abnormal Criteria for Vital Signs

Variable	Unit	Low	High
SBP	mmHg	≤ 90 mmHg	> 150 mmHg
DBP	mmHg	≤ 50 mmHg	> 90 mmHg
Heart rate	bpm	< 40 bpm < 50 bpm < 50 bpm and decrease from predose (Baseline) of > 10 bpm at 4 hours on Day 1 or remonitoring visit	>100 bpm

18.7. PHYSICAL EXAMINATION

A listing of all physical examination assessments in subjects who had at least 1 abnormal physical examination finding will be provided.

18.8. OTHER SAFETY ASSESSMENTS

18.8.1. PULMONARY FUNCTION TESTS

The following pulmonary function test (PFT) measurements (actual and % Predicted) will be reported for this study:

- Forced Expiratory Volume at 1 second (FEV1)
- Forced Vital Capacity (FVC)
- Total Lung Capacity (TLC)
- FEV1/FVC
- Forced Expiratory Flow (FEF) 25-75
- Diffusing Capacity of the Lungs for Carbon Monoxide (DLCO) (if available)

The following summaries will be provided for PFT data:

- Value and change from Baseline by visit
- Incidence of markedly abnormal values by visit

All PFT data will be listed. A listing in subjects who ever reported an abnormality in PFT will also be provided, including a flag for whether markedly abnormal criterion is also met.

18.8.2. PFT Markedly Abnormal Criteria

Markedly abnormal quantitative PFT measurements will be identified in accordance with the following predefined markedly abnormal criteria:

- % Predicted FEV1 < 50%
- % Predicted FVC < 50%
- % Predicted FEV1/FVC ratio < 50%

Potentially important PFT measurements will also be identified using the criteria below:

- Decrease from Baseline > 20% in FEV1, ie, percent change from Baseline < -20%
- Decrease from Baseline > 20% in FVC, ie, percent change from Baseline < -20%
- Decrease from Baseline > 20% in DLCO, ie, percent change from Baseline < -20%

18.8.3. Ophthalmoscopy and Optical Coherence Tomography (OCT)

The following summaries will be provided for ophthalmoscopy and OCT data:

- Values and change from Baseline in central foveal thickness by visit
- Categorical result in ophthalmoscopy with OCT parameters by visit (the categories are listed on the eCRF)
- Only PFT and OCT assessments completed by 50 or more subjects in the SAF will be presented in the summaries.

A listing of all ophthalmoscopy and OCT assessments will be provided. A listing of subjects who ever reported an abnormality in OCT will also be provided.

19. Pharmacokinetics

Plasma concentrations of etrasimod, and if warranted M3 (AR503641) and M6 (AR504344), will be assessed from samples collected prior to dosing and 4 hours (± 15 minutes) post-dose (after ECG) on Week 0/Day 1, samples collected prior to dosing (trough) at Weeks 4, 8, 16, 24 of the double-blind period, and Weeks 28, 32, 46, 52 or the ET visit, if applicable, of the extension period. Additionally, a 4-Week Safety Follow-up visit sample is being collected. Concentrations below the limit of quantification (BLQ) will be assigned a numerical value of zero for the calculation of descriptive statistics in concentrations (except for geometric mean and geometric %CV) and plotting of concentrations. For geometric mean and geometric %CV, the zero values will be excluded.

The pharmacokineticist will determine the strategy for dealing with data affected by protocol deviations or events which may impact the quality of PK concentration data on a case-by-case basis with input from the Pfizer study physician and Pfizer clinical pharmacologist, as needed. Examples for protocol deviations or events include, but may not be limited to, vomiting on the day prior to trough sample collection, sample processing errors that lead to inaccurate bioanalytical results, and/or inaccurate dosing prior to PK sampling. In the case of an important protocol deviation or event, the affected PK data collected may be excluded from the summaries, calculation of the average steady state trough plasma concentration (C_{trough}) based on Week 4 to Week 24 trough concentrations ($C_{trough,ss,W4W24}$), exposure-response analysis, and/or population PK analysis, but will still be reported in the study result listings.

Unless otherwise specified, PK summaries will use the Pharmacokinetic Set. Exposure-response analysis will use the Safety Set.

Individual subject etrasimod plasma concentrations (and etrasimod metabolites, if applicable) will be presented in the data listings, including subject ID, treatment received, sex, age, weight, nominal timepoint, actual blood collection date/time, concentration, and time since last dose, and also summarized using descriptive statistics (n, mean, SD, %CV, geometric mean, geometric % CV, median, minimum, and maximum) by nominal timepoint and treatment.

All subjects are expected to achieve steady-state plasma concentration by Week 2. For each subject, the average steady state trough (pre-dose) plasma concentration (C_{trough}) based on Week 4 to Week 24 trough concentrations ($C_{trough,ss,W4-W24}$) per subject will also be calculated and presented in a data listing and summarized using descriptive statistics by treatment.

Mean (±SD) etrasimod concentration versus nominal time will be plotted by treatment on linear scale for the double-blind period and the extension period separately. Additionally, a mean (±SD) concentration-time plot with etrasimod and metabolites overlaid on linear scale will be generated for each treatment for the induction period and the extension period separately.

Box plots of etrasimod, M3, and M6 concentrations at Day 1, 4 hour post-dose timepoint (C4hr) and steady-state C_{trough} (Weeks 2 through 24 [$C_{trough,ss,W4-W24}$]) will be plotted versus nominal timepoint on ordinal scale, separately by treatment. Similarly, box plots of etrasimod, M3, and M6 concentrations will be presented for the extension phase by treatment.

Individual subject plasma concentrations versus actual time will be plotted since start of treatment on linear scale for the induction period (Weeks 0 to 24) and the extension period (Weeks 28 to 52) separately. Individual subject plots with etrasimod and metabolite concentration data will be overlaid on the same plot.

Scatter plots of individual subject average steady-state $C_{trough\ ss,W4-W24}$ versus Baseline body weight and versus Baseline age as continuous variables will be generated by treatment for the Pharmacokinetic Set, which will also include the Spearman's rank correlation coefficient, p-value, and LOESS trend line.

To explore potential etrasimod plasma exposure-response relationships, the following scatter plots will be generated in the Safety Set, including data from placebo subjects (plotted at 'zero' concentration). Any contributing data from the placebo subjects will be annotated with a different color or symbol. All scatter plots will include the Spearman's rank correlation coefficient, p-value, and LOESS trend line (based on subjects that received active treatment only).

- Individual subject percentage change in esophageal PEC versus etrasimod steady-state C_{trough} at Week 16
- Individual subject lymphocytes versus etrasimod steady-state C_{trough} by visit (Weeks 4 through 24)
- Individual subject absolute changes from Baseline in lymphocytes versus etrasimod steady-state C_{trough} by visit (Weeks 4 through 24)
- Individual subject percent changes from Baseline in lymphocytes versus etrasimod steady-state Ctrough by visit (Weeks 4 through 24)
- Individual subject heart rate versus etrasimod concentration at matching visit/timepoint (Day 1 through Week 24): two-panel plot with Day 1 on the left and Weeks 4 through 24 on the right
- Individual subject absolute changes from Day 1 Pre-dose in heart rate versus etrasimod concentration at matching visit/timepoint (Day 1 through Week 24): two-panel plot with Day 1 on the left and Weeks 4 through 24 on the right

Furthermore, the following overlay plots will be generated in the Pharmacokinetic Set within the etrasimod group only:

- Mean (SD) absolute change from Baseline in lymphocytes and mean (SD) etrasimod concentration versus nominal timepoint (Weeks 0-24)
- Mean (SD) percent change from Baseline in lymphocytes and mean (SD) etrasimod concentration versus nominal timepoint (Weeks 0-24)
- Mean (SD) heart rate and mean (SD) etrasimod concentration versus nominal timepoint (Weeks 0-24)
- Mean (SD) absolute change from Baseline in heart rate and mean (SD) etrasimod concentration versus nominal timepoint (Weeks 0-24)

All summary figures, with the exception of the etrasimod and metabolite overlay plots, scatter plots of individual subject $C_{trough,ss,W4-W24}$ values versus Baseline body weight and versus Baseline age, scatter plots of individual subject heart rate versus etrasimod concentration at matching visit/timepoint (Day 1 through Week 24), and scatter plots of individual subject absolute changes from Day 1 Pre-dose in heart rate versus etrasimod concentration at matching visit/timepoint will be repeated by:

- Sex
- Weight (≤ Median or > Median)
- Age at consent
- Tobacco use (Yes/No)

The plasma concentrations over time will be used in a population PK analysis, which will be described in a separate plan.

20. DATA NOT SUMMARIZED OR PRESENTED

The other variables and/or domains not summarized or presented are:

- Comments

These domains and/or variables will not be summarized or presented, but will be available in the clinical study database, SDTM and/or ADaM datasets.

21. REFERENCES

ICH E3. Structure and Content of Clinical Study Reports. Dated 30 November 1995. Available at: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf (accessed 14 August 2018).

ICH E9. Statistical Principles For Clinical Trials. Dated 5 February 1995. Available at: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf (accessed 14 August 2018).

ICH E9 R(1). Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials. Dated 20 November 2019. Available at: https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf.

Rubin, D.B. (1987). Multiple Imputations for Nonresponse in Surveys. Wiley.

FDA. Statistical Considerations for Clinical Trials During the COVID 19 Public Health Emergency: Guidance for Industry. Center for Drug Evaluation and Research, Food and Drug Administration. https://www.fda.gov/media/139145/download. Published 2020. Accessed February 22, 2021.

APPENDICES

APPENDIX 1. PROGRAMMING CONVENTIONS FOR OUTPUTS

Output Conventions

Outputs will be presented as shown in the Output shells.

Decimals, Percentages, and P-Values

- If the original data has N decimal places, then the summary statistics should have the following decimal places:
- Minimum and maximum: N
- Mean (and LS Means), median (and Q1, Q3): N + 1
- SD or SE: N + 2
- Percentages will be reported to one decimal place. Where counts are zero, percentages will not appear in the output.
- P-values will be reported to four decimal places, except values <1.000 but >0.999 will be presented as '>0.999' (e.g., 0.9998 is presented as >0.999); and
- values <0.0001 will be presented as '<0.0001' (e.g., 0.00009 is presented as <0.0001). Rounding will be applied after the <0.0001 and >0.999 rule.

Dates & Times

Depending on data available, dates and times will take the form DDMMMYYYY or DDMMMYYYY:hh:mm.

Spelling Format

English US.

Presentation of Treatment Groups

For outputs, treatment groups will be represented as follows and in that order:

Treatment Group	For Tables and/or Graphs	For Listings
etrasimod 2 mg	etrasimod 2 mg	etrasimod 2 mg
etrasimod 1mg	etrasimod 1 mg	etrasimod 1 mg
etrasimod any dose*	etrasimod any dose	Placebo
Placebo	Placebo	Screen Failure
Screen Failure	Screen Failure	

Treatment Group	For Tables and/or Graphs	For Listings
Not Treated**	Not Treated	Placebo

* To be used for tables and/or graphs only.

** To be used for subjects in safety listings who are randomized but do not receive study drug.

Presentation of Visits

For outputs, visits will be represented as follows and in that order when applicable:

Visit Name*	Study Period
Screening	Screening
Baseline, Day 1, Week 2, Week 4, Week 8, Week 16, Week 20, Week 24	Double-Blind Treatment
Week 28, Week 32, Week 46, Week 52	Extension Treatment
2-Week Follow-Up, 4-Week Follow-Up	Safety Follow-Up

- *If a visit is not available for a certain analysis, it won't be displayed

LISTINGS

All listings will be ordered by the following (unless otherwise indicated in the template):

- Randomized treatment group (or treatment received if it's a safety output), first by Etrasimod 2 mg, Etrasimod 1 mg, then Placebo, then Screen Failure and then No Treatment (only in safety listings if there are any randomized subjects who did not receive study drug)
- Subject number (which is expected to incorporate study site/center),
- Date (where applicable)

APPENDIX 2. PARTIAL DATE CONVENTIONS

Imputed dates will NOT be presented in the listings.

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS:

AE START DATE	AE STOP DATE	ACTION
Known	Known, Partial or Missing	If AE start date < study drug first dose date, then not TEAE If AE start date ⩾ study drug first dose date, then TEAE
Partial, but known components show that it cannot be on or after date of first dose of study drug	Known, Partial or Missing	Not TEAE
Partial, could be on or after date of first dose of study drug	Known	If AE stop date < study drug first dose date, then not TEAE If AE stop date ⩾ study drug first dose date, then TEAE
	Partial	Impute AE stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If AE stop date < study drug first dose date, then not TEAE If AE stop date ⩾ study drug first dose date, then TEAE
	Missing	Assumed TEAE
Missing	Known	If AE stop date < study drug first dose date, then not TEAE If AE stop date ⩾ study drug first dose date, then TEAE
	Partial	Impute AE stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If AE stop date < study drug first dose date, then not TEAE If AE stop date ⩾ study drug first dose date, then TEAE

AE START DATE	AE STOP DATE	ACTION
	Missing	Assumed TEAE

ALGORITHM FOR PRIOR AND CONCOMITANT MEDICATIONS:

START DATE	STOP DATE	ACTION
Known	Known	If medication stop date < study med first dose date, assign as prior If medication stop date >= study med first dose date, assign as concomitant
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If medication stop date < study med first dose date, assign as prior If medication stop date ⩾ study med first dose date, assign as concomitant
	Missing	If medication stop date is missing could never be assumed a prior medication, assign as concomitant
Partial	Known	Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown), then: If medication stop date < study med first dose date, assign as prior If medication stop date ⩾ study med first dose date, assign as concomitant
	Partial	Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown) and impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If medication stop date < study med first dose date, assign as prior If medication stop date ⩾ study med first dose date, assign as concomitant
	Missing	Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown), then: If medication stop date is missing could never be assumed a prior medication, assign as concomitant

START DATE	STOP DATE	ACTION
Missing	Known	If medication stop date < study med first dose date, assign as prior Else assign as concomitant
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If medication stop date < study med first dose date, assign as prior If medication stop date ⩾ study med first dose date, assign as concomitant
	Missing	Assign as concomitant

APPENDIX 3. QUESTIONNAIRE SCORING INFORMATION

DSQ

The DSQ is a validated patient-reported outcome questionnaire that will be used to characterize baseline dysphagia severity and changes with treatment intervention. Questions 1 and 2 of the DSQ will be used to determine the eligibility regarding active dysphagia during the Screening Period. One episode of dysphagia is counted when a subject responds "yes" to questions 1 and 2. Subjects must have at least an average of 2 episodes of dysphagia per week over 2 weeks during the Screening Period.

The DSQ will be scored over a 14-day period, and a minimum of 8 DSQ daily subject eDiary entries out of the 14-day period preceding the date of scoring are required to calculate DSQ score. The DSQ total score range is 0 to 84.

PGIC

The PGIC is a 1-item questionnaire using a 5-point Likert scale to capture the subject's overall impression of change in EoE symptoms from Baseline. The recall period is 16 to 52 weeks. The PGIC score range is 1 to 5.

PGIS

The PGIS is a current-state, 1-item questionnaire using a 4-point Likert scale to capture the subject's overall impression of EoE symptom severity over the past 7 days. The PGIS score range is 1 to 4.

EoE-HSS

The EoE-HSS objectively assesses histologic changes in esophageal mucosa beyond eosinophil number. It describes changes in 8 histologic features and assesses severity (grade) and extent (stage) of abnormalities that are scored using a 4-point scale (0 = normal; 3 = most severe or extensive). The 8 features include:

1. Eosinophilic inflammation
2. Basal zone hyperplasia
3. Eosinophil abscess
4. Eosinophil surface layering
5. Dilated intercellular spaces
6. Surface epithelial alteration
7. Dyskeratotic epithelial cells
8. Lamina propria fibrosis

EREFS

The EREFS will be used to characterize the baseline endoscopic severity of EoE and changes with treatment. Presence and severity of 6 esophageal endoscopic features are scored by the local endoscopist. The endoscopic features include:

1. Fixed Rings
2. Exudates
3. Furrows
4. Edema
5. Stricture
6. Crepe paper esophagus

EoE-QOL-A

The EoE-QOL-A 37-item questionnaire is a disease-specific measure of health-related QOL in adults with EoE. It is scored using a 5-point scale and measures 5 domains:

1. Eating/Dietary Impact
2. Social Impact
3. Emotional Impact
4. Disease Anxiety
5. Choking Anxiety

APPENDIX 4. IDENTIFICATION OF RESCUE THERAPY

This appendix outlines the process the Pfizer clinical team/medical reviewers will use to determine rescue therapy for EoE in the APD334-206 study.

The objective of rescue therapy identification is to establish 1) intercurrent events for the efficacy estimands, and 2) exclusion of subjects from the Per Protocol Set(s).

Per protocol, rescue therapy is defined as any new therapy (ie, medication or procedure) or change in existing EoE therapy used to treat new or worsened EoE symptoms. In addition, use of certain medications/procedures that do not meet the definition of rescue therapy but are deemed to have significant impact on the efficacy assessment will be reviewed and may be considered as rescue therapy.

Only therapies reported on the electronic case report forms (eCRFs) will be assessed for rescue therapy determination. Use of rescue therapy during the follow-up period (beginning on or after the date of last study treatment administration) will not be considered as a rescue therapy.

The list of all rescue therapy uses in the study will be identified and finalized prior to study unblinding.

Medical procedure

- Esophagogastroduodenoscopy (EGD)
  - Rule:
    - When dilation was done during EGD, any after first dose
    - When dilation was not done during EGD, any EGD with food bolus removal after first dose

Medications

- Proton Pump Inhibitors
  - Examples of Medications
    - Omeprazole (Prilosec)
    - Lansoprazole (Prevacid)
    - Dexlansoprazole (Dexilant)
    - Esomeprazole (Nexium)
    - Pantoprazole (Protonix)
    - Rabeprazole (Aciphex)
  - Rules
    - When used for EoE: any new initiation or change in dose after Day 1
    - When used for other disease (eg, GERD): New initiation or change in dose after Day 1 AND
      - Use > 4 weeks at the time of EGD OR
      - Use > 8 weeks with the end of treatment date falling within 4 weeks of EGD
- Corticosteroids
  - Applicable route of administration: swallowed, oral, IM, SC, IV
  - Examples of Medications
    - Fluticasone (Flovent)
    - Budesonide (Pulmicort, Flexhaler)
    - Betamethasone
    - Cortisone
    - Deflazacort
    - Dexamethasone
    - Fludrocortisone
    - Hydrocortisone
    - Methylprednisolone
    - Predinisone
    - Prednisolone

        - Triamcinolone
    - Rules
        - When used for EoE: any new initiation after Day 1
        - When used for other diseases: any new initiation after Day 1 AND
            - Use > 7 days at the time of EGD OR
            - Use > 14 days with the end of treatment date falling within 8 weeks of EGD

- Drugs that have been or are being investigated in clinical trials for the treatment of EoE
    - Examples of medications
        - Anti-interleukin (IL)-5 antibodies (eg, mepolizumab, reslizumab, benralizumab)
        - IL-4 and IL-13 antagonists (dupilumab)
        - Anti-IgE antibodies (omalizumab)
    - Rule
        - Any new initiation after Day 1
- Other protocol-prohibited immunomodulatory agents
    - Examples of medications: please refer to the protocol section 6.7.3. Prohibited Concomitant Therapy
    - Rule
        - Any new initiation after Day 1

Electronic Signature Manifestation

This page is a manifestation of the electronic signature(s) used in compliance with the organization's electronic signature policies and procedures.

Signer Full Name	Meaning of Signature	Date and Time
PPD	Document Approval (I certify that I have the education, training and experience to perform this task)	13 Jan 2023 15:28:28 UTC
PPD	Document Approval (I certify that I have the education, training and experience to perform this task)	13 Jan 2023 16:17:59 UTC
PPD	Document Approval (I certify that I have the education, training and experience to perform this task)	13 Jan 2023 18:19:56 UTC