Statistical Analysis Plan DEGAS 7810

Title Page

Study Title	A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED CLINICAL TRIAL TO EVALUATE THE EFFICACY OF DUPILUMAB (ANTI-IL4RA) IN SUBJECTS WITH EOSINOPHILIC GASTRITIS
Short Title	Dupilumab Eosinophilic Gastritis Study (DEGAS)
Study Registration Number	NCT03678545
Sponsor	Marc Rothenberg MD, PhD Cincinnati Children's Hospital Medical Center Cincinnati OH 45229
Intervention	DUPILUMAB
Investigation Agent Manufacture	Regeneron Pharmaceuticals
Protocol Number	V5.0
Phase	Phase II
Funding Mechanism	U54AI117804
Statistical Analysis Plan Version	V2.4
Statistical Analysis Plan Date	February 12, 2025
Author	Xue Zhang, PhD Cincinnati Children's Hospital Medical Center 3333 Burnet Av Cincinnati, OH, 45229

Statistical Analysis Plan Approval Signature Page

AUTHOR: Name: Xue Zhang, PhD
Position/Title: Biostatistician
Signature and Date: _____ 2/12/2025

APPROVED BY: Name: Lisa J. Martin
Position/Title: Faculty Biostatistician
Signature and Date: _____ 2/12/2025

SAP Revision History

Version	Date	Changes
1.0	9/19/2023	Development of initial SAP
2.0	6/18/2024	1. Placing SAP on standardized template 2. Adding variable definitions for derived variables. 3. Changing EGID abbreviations to updated standard
2.1	7/17/2024	Added hierarchical testing structure for key secondary outcomes
2.2	8/1/2024	Changed the rules for the window around the symptom diary reporting for EOT. Using time since baseline may have impact on missing rate.
2.3	12/5/2024	Change to remove internal inconsistency. When hierarchical testing structure was added EoG HSS was not moved from an exploratory outcome to a key secondary outcome. Based on findings from the clinical trial of benralizumab in EoG, the study team agreed that EoG HSS was a key secondary outcome. When the study was initially designed, EoG HSS had not been evaluated clinically, thus it was not placed as a key secondary outcome.
2.4	2/12/2025	Changed the analysis population for the key primary and secondary outcomes to the complete case ITT due to similarity of results using imputed data and limited missingness.

SAP Roles

SAP Primary Author: Xue Zhang
Senior Statistician: Lisa J. Martin
SAP Contributor – SAP Reviewer: Kara Kliewer
SAP Contributor – Project Manager: Regina Yearout

List of Abbreviations

ACT	Asthma Control Test
AE	Adverse Event
AEC	Absolute Eosinophil Count
BP	Blood Pressure
CBC	Complete Blood Count
CCHMC	Cincinnati Children's Hospital Medical Center
CEGIR	Consortium of Eosinophilic Gastrointestinal Disease Researchers
CFR	Code of Federal Regulations
CGI-C	Clinician Global Impression of Change
CGI-S	Clinician Global Impression of Severity
CRF	Case Report Form
COVID	Coronavirus Disease
CTCAE	Common Terminology Criteria for Adverse Events
DAIT	Division of Allergy, Immunology, and Transplantation
DBP	Double Blind Phase
DEGAS	Dupilumab Eosinophilic Gastritis Study
DNA	Deoxyribonucleic Acid
DSMB	Data Safety Monitoring Board
EoD	Eosinophilic Duodenitis
EoE	Eosinophilic Esophagitis
EoG	Eosinophilic Gastritis
EGD	Esophagogastroduodenoscopy
EGID	Eosinophilic Gastrointestinal Disorder
EoG-EREFS	Eosinophilic Gastritis Endoscopy Assessment
EoG-SQ	Eosinophilic Gastritis Symptom Questionnaire
EOS	End of Study (Completion of telephone contact 12-weeks after last dose of study drug)
EOT	End of Treatment
EREFS	Endoscopic Reference Score
FDA	Food and Drug Administration
GCP	Good Clinical Practice
hpf	High Power Field
HSS	Histology Scoring System
ICF	Informed Consent Form
ICH	International Conference on Harmonization
IgE	Immunoglobulin E
IL	Interleukin
IND	Investigational New Drug
IPM	Independent Protocol Monitor
IRB	Institutional Review Board
IP	Investigational Product
ITT	Intention to Treat
mAb	Monoclonal Antibody
MCAR	Missing at Completely Random
MO	Medical Officer
MOP	Manual of Procedures
NIAID	National Institute of Allergy and Infectious Diseases
OLE	Open Label Extension
PP	Per Protocol
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity

PHI	Protected Health Information
PI	[Site] Principal Investigator
PM	Project Manager
PoR	Pharmacist of Record
PRO	Patient Reported Outcome
PROMIS	Patient Reported Outcomes Measurement Information System
QC	Quality Control
QOL	Quality of Life
RDCRN	Rare Diseases Clinical Research Network
RNA	Ribonucleic Acid
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAR	Suspected Adverse Reaction
SC	Subcutaneous
SF-12	Short Form 12 Health Survey
SOA	Schedule of Activities
SODA	Severity of Dyspepsia Assessment
SOP	Standard Operating Procedure
SUSAR	Serious Unexpected Suspected Adverse Reaction
TEAE	Treatment-Emergent Adverse Event
Th2	Type 2 Helper

Introduction

The primary objective of this study is to assess the efficacy and safety of Dupilumab in the treatment of Eosinophilic Gastritis. This document outlines the statistical methods to be implemented during the analyses of data collected within the scope of Consortium of Eosinophilic Gastrointestinal Disease Researchers' Protocol CEIGR 7810 titled "A Randomized, double-blind, placebo-controlled clinical trial to evaluate the efficacy of dupilumab (anti-IL4rα) in subjects with eosinophilic gastritis". This Statistical Analysis Plan (SAP) was prepared in accordance with CEGIR 7810 protocol, V6.0, dated 12/6/2023.

The purpose of this Statistical Analysis Plan (SAP) is to provide a framework in which answers to the protocol objectives may be achieved in a statistically rigorous fashion, without bias or analytical deficiencies, following methods identified prior to database lock. Specifically, this plan has the following purposes:

- To prospectively outline the specific types of analyses and presentations of data that will form the basis for conclusions.
- To explain in detail how the data will be handled and analyzed, adhering to commonly accepted standards and practices of biostatistical analysis. Any deviations from these guidelines must be substantiated by sound statistical reasoning and documented in writing in the final clinical study report (CSR) or manuscript.

The analyses described in this analysis plan are consistent with the analyses described in the study protocol. The order may be changed for clarity. If there are discrepancies between the protocol and SAP, the SAP will serve as the definitive analysis plan.

Any analysis performed not prospectively defined in this document will be labeled as post hoc & exploratory.

Overview and Objectives of Study Design

Overview: This is a phase 2, multi-center, randomized, double-blind, placebo-controlled trial testing the efficacy of dupilumab vs. placebo in eosinophilic gastritis (EoG) with or without concurrent esophageal and/or duodenal eosinophilia. Potential participants will be screened during an 8-week screening period. An esophagogastroduodenoscopy (EGD) with biopsies will be performed to determine study eligibility, and participants will be enrolled based on the presence of active disease and their ability to meet the study inclusion and exclusion criteria. Qualifying participants will be blindly randomized 1:1 to either study drug(dupilumab) or placebo and will receive 600 mg once followed by 300 mg doses every two weeks of study treatment (for a total of 6 injections). During the treatment period, participants will be monitored for adverse events/reactions and will complete patient reported outcome metrics to track their symptoms and general wellbeing. At the end of the 12-week randomized treatment period (EOT), a repeat endoscopy with biopsies will be performed to assess the change in gastric eosinophil count from pre-to-post treatment. Participants and site study staff will remain blinded to the endoscopy biopsy results until all participants have completed the DBP EOT visit. Study participants who complete the 12-week treatment phase (whether on drug or placebo) will have the option to continue into an open-label extension (OLE) and laboratory studies and patient reported outcomes will be completed as indicated in the schedule of activities (SOA). The OLE ensures that every participant gets access to the drug and will provide data on the effect of a longer treatment period in comparison to the initial 12 weeks in the double-blind phase. A final study EGD will be performed two weeks after the last dose of the study drug. All participants will be followed for an additional 12 weeks after the last dose of study drug regardless of whether they participate in the OLE.

Objectives: The central hypothesis of this study is that dupilumab is safe and efficacious for reducing eosinophilia and a Th2 gene profile in EoG.

Primary Objective for the Double-Blind Phase (DBP) Population

To assess the efficacy of repeat subcutaneous (SC) doses of dupilumab, compared with placebo, to reduce eosinophilic inflammation in the gastrointestinal tract of participants with EoG with or without concurrent esophageal and/or duodenal eosinophilia involvement.

Secondary Objectives for the DBP Population

- To assess the efficacy of dupilumab compared to placebo as measured by:
  - Patient global impression of severity (PGI-S)
  - Patient global impression of change (PGI-C)
  - Proportion of participants achieving threshold response < 30 eos/hpf in 5 hpfs
  - Eosinophilic Gastritis Endoscopy Assessment (EoG-EREFS)
  - Eosinophilic Gastritis Symptom Questionnaire (EoG-SQ)
  - Comprehensive assessment of histologic features (EoG HSS)

Exploratory Objectives for the DBP Population

- To assess the efficacy of dupilumab compared to placebo as measured by:
  - Severity of Dyspepsia Assessment tool (SODA)
  - Clinician global impression of severity (CGI-S)
  - Clinician global impression of change (CGI-C)
  - Quality of Life Questionnaire (SF12, PROMIS 29, PROMIS 25)
  - EoE Endoscopic Reference Score (EREFS) (participants with esophageal involvement)
  - Blood and tissue biomarkers and gastric transcripts
  - Peripheral eosinophil counts, hemoglobin & hematocrit, total protein and albumin, total IgE, iron storage panel (iron, ferritin, total iron-binding capacity)
  - Eosinophilic Duodenitis- Endoscopic Reference Score (EoD-EREFS)

Exploratory Objectives for the OLE Population

Key Objective:

To assess the efficacy of repeat SC doses of dupilumab to reduce eosinophilic inflammation in the gastrointestinal tract of participants with EoG with or without esophageal and/or duodenal eosinophilia.

- To assess the efficacy of dupilumab as measured by:
  - Patient global impression of severity (PGI-S)
  - Patient global impression of change (PGI-C)
  - Proportion of participants achieving threshold response < 30 eos/hpf in 5 hpfs
  - Eosinophilic Gastritis Endoscopy Assessment (EoG-EREFS)
  - Eosinophilic Gastritis Symptom Questionnaire (EoG-SQ)
  - Comprehensive assessment of histologic features (EoG HSS)

- To assess the efficacy of dupilumab as measured by:
  - Severity of Dyspepsia Assessment (SODA)
  - Clinician global impression of severity (CGI-S)
  - Clinician global impression of change (CGI-C)
  - Quality of Life Questionnaire (SF12, PROMIS 29, PROMIS 25)
  - EoE Endoscopic Reference Score (EREFS) (participants with esophageal involvement)
  - Blood and tissue biomarkers and gastric transcripts
  - Peripheral eosinophil counts, hemoglobin & hematocrit, total protein and albumin, total IgE, iron storage panel (iron, ferritin, total iron-binding capacity)
  - Eosinophilic Duodenitis- Endoscopic Reference Score (EoD-EREFS)

Safety objectives

- Incidence, occurrence, and severity of AEs/SAEs
- Systemic tolerance (safety laboratory assessments)

Sample Size Considerations

As we recognize that not all participants may complete the study, a dropout rate of 10% is assumed for all sample size calculations provided below.

It is assumed that the true mean percent reduction from baseline in mean gastric eosinophil counts when treated with dupilumab will be at least 50%. This is a conservative estimate based on the efficacy of dupilumab in EoE patients; dupilumab reduced the mean (least square) peak esophageal eosinophil count by -94/HPF (9.5 standard error; n=23) and -7.4 (9.6 standard error; n=24) in the placebo group in a recent study of EoE13. At Week 12, dupilumab versus placebo significantly reduced peak eosinophil counts from baseline (−92.9% vs +14.2%; P<0.0001). It is also assumed that the true mean percent reduction from baseline in mean eosinophil counts for the placebo group will be at most 25% and we expect it to be much less as the placebo effect in EoE was 14% increase in eosinophil levels as noted above. Assuming a standard deviation of 25 percentage units, a sample size of 20 participants per treatment group will provide at least 80% power to detect a true absolute difference in percent reduction from baseline in mean eosinophil counts of 23% at alpha = 0.05. If we assume that the dropouts from the dupilumab arm have a mean reduction in eosinophil counts similar to those on placebo (25%) then we would expect 48% decline in eosinophil count among those in the dupilumab arm, thus we are sufficiently powered even with dropouts. A sample size of 20 per treatment group will also provide at least 80% power to detect a difference in the proportions of participants with induction of disease remission of 42 percentile points (e.g. 62% vs. 20%) at the two-sided 5% level of significance. Notably, if 2 individuals in the dupilumab arm drop out then the effective disease remission rate would be 67% among those completing the study in order to obtain a 62% overall remission rate for the dupilumab arm.

Randomization, stratification, and blinding

Randomization will be accomplished via a central variable block randomization scheme provided by an independent statistician to the study pharmacist (or qualified designee). Block sizes will vary between 2 and 4. In addition, the randomization will be stratified based on age (≥ 12 to < 18 and ≥18 years) and whether the patient uses or does not use systemic corticosteroids, (e.g. prednisone at daily dosage of 5 mg or higher), swallowed glucocorticosteroids for EoG (Entocort), or other systemic immunosuppression therapy at the time of randomization. The randomization schema will ensure that potential confounding factors are equally distributed among the placebo and drug treatment groups and will eliminate ‘researcher selection bias’.

The investigators, site staff other than the unblinded site pharmacist/pharmacy staff, pathologists, and study participants will all remain blinded to treatment assignment and IgE levels throughout the double blind (IgE will be analyzed after participants have completed the study). The IPM, DAIT/NIAID MO, DAIT/NIAID PM, and any other personnel designated by the sponsor (Dr. Rothenberg as PI of the study) who have regular contact with the study site will remain blinded to all treatment assignments. Unblinding of the data will occur after initial data quality review to ensure data queries for key outcome measures.

Definitions of Patient Populations To Be Analyzed

Patient drop out is always a concern for clinical trials as if the participants who drop out are not missing at completely random (MCAR), the results may be biased. Based on our prior study of diet and eosinophilic esophagitis, a related eosinophilic condition of the gastrointestinal tract, we expect a low rate of participant

drop out (3.9%). Missing data below 5% is considered negligible, and many studies use only the complete data [26]. However, our analytic strategy will be to compare all randomized participants who receive study drug (complete V1; intent to treat sample) in the primary analyses using multiple imputation to complete case. If the results yield similar results, the complete case will be the primary report, with the imputed results in supplemental data. The per protocol population analyses will be conducted as secondary (sensitivity) analyses for all outcomes. Participants who complete the blinded phase and who do not have any major protocol deviations will be considered in the per protocol set. Examples (not all inclusive) of major protocol deviations include missing more than one dose of IP or steroid use within 30 days of endoscopy.

The primary analyses will be intention-to-treat (all participants who receive at least one dose of study drug will be included), and the per protocol population analyses will be conducted as secondary (sensitivity) analyses for all outcomes. Participants who complete the blinded phase and who do not have any major protocol deviations will be considered in the per protocol set. Examples (not all inclusive) of major protocol deviations include missing more than one dose of IP or steroid use within 30 days of endoscopy.

For the intent to treat model, it is also possible that some participants may be exposed to extended dosing due to an inability to perform the endoscopy as scheduled due to a public health emergency or illness precluding undergoing an endoscopy on schedule. If a subset of participants undergo extended dosing, we will compare individuals on study drug who had extended dosing to those who did not (sensitivity analysis) to see if the primary study outcome was substantially different based on extended dosing and will perform analyses only on those who did not undergo extended dosing to ensure results are consistent.

Endpoints

Primary Endpoint

The primary outcome measure will be the relative change in the mean of the eosinophil counts from the 5 worst microscopic fields (most eosinophil dense hpfs) in the gastric antrum and/or body. Relative change is defined as (week 12 – baseline)/baseline. Regression analyses will be performed with the primary outcome measure as the dependent variable and treatment group as the independent variable. Appropriate transformations, such as a log-transformation, may be used to satisfy the distributional assumptions prior to model fitting. If there are continued concerns about data distributions, equivalent nonparametric tests will be employed. To ensure that immunosuppression (e.g. corticosteroid, swallowed glucocorticosteroids for EoG (Entocort), or immunosuppressive agents) use does not influence results, we will also perform regression analyses with corticosteroid/immunosuppression therapy as a covariate (including considering binary (y/n), and dose (none, low, high).

Secondary Endpoints

Secondary endpoints include
1) Absolute change in the mean of the peak eosinophil counts in the 5 most eosinophil dense HPFs in the gastric antrum and/or body.
2) Absolute change in PGI-S score.
3) PGI-C score at 12 weeks.
4) Proportion of participants achieving threshold response < 30 eos/hpf in all 5 hpfs.
5) Absolute change in endoscopic score as measured by EoG-EREFS.
6) Absolute change in the total EoG-SQ eDiary score.
7) Absolute change in proportion of days with any (1 or more) symptom as measured by the EoG-SQ eDiary.
8) Absolute change in histologic features measured by EoG HSS

Hierarchical Testing structure

We will employ a hierarchical testing strategy to minimize the impact of multiple testing for the key secondary outcomes from the double blind. Thus, p-values will be reported only when prior outcomes had reached

significance ($p \leq 0.05$). Once an outcome fails to reach significance the confidence intervals of the estimates will be reported.

	End Points	Order
Primary	Relative change in mean of the peak gastric count	1
Key Secondary	Absolute change in EoG-HSS	2
	Absolute change in mean of the peak gastric count	3
	Absolute change in EoG-REFS (total score – sum of antrum, body, fundus)	4
	Absolute change in the total EG-SQ eDiary score	5
	Proportion achieving response (< 30 eos/hpf in all 5 hpfs)	6
	Absolute change in PGI-S score	7
	PGI-C score	8
	Absolute change in proportion of days of any (1 or more) symptom in EoG-SQ	9

Exploratory Endpoints

Double-blind

1) Absolute change in SODA, SF-12, PROMIS 29, PROMIS 25, and CGI-S scores
2) CGI-C score at 12 weeks
3) Absolute change in proportion of days with each individual symptom and multiple symptoms measured by the EoG-SQ eDiary.
4) Absolute change in each individual EoG-SQ eDiary symptom score
5) Induction of disease remission (percentage who achieve histological remission in the stomach as defined by peak eosinophil counts less than 30/hpf in 1, 2, 3 and 4 of the 5 hpfs identified for the primary endpoint).
6) For participants with esophageal involvement and/or duodenal eosinophilia involvement, absolute change in EoE HSS and EoD HSS respectively.
7) Absolute and relative change in blood and tissue biomarkers and gastric transcripts
8) Absolute and relative change in blood eosinophil count, hemoglobin, hematocrit, total protein levels, albumin, IgE, iron storage panel (iron, ferritin, total iron-binding capacity)
9) The association of blood and tissue biomarkers and transcripts with drug responsiveness.
10) Absolute changes in esophageal and duodenal peak eosinophil levels.
11) Absolute changes in EoE-EREFS and EoD-EREFS in participants with esophageal and/or duodenal involvement, respectively.

For double-blind endpoints above, a comparison between drug vs placebo at 12 weeks will be made.

Open Label Extension

Two time points will be evaluated for the OLE: baseline to OLE/EOT and DBP/EOT to OLE/EOT for the endpoints listed below:

Key Endpoint/Outcome
Relative change of the mean of the peak eosinophil counts in the 5 most eosinophil dense HPFs in the gastric antrum and/or body.

1) Absolute change in the mean of the peak eosinophil counts in the 5 most eosinophil dense HPFs in the gastric antrum and/or body
2) Absolute change in PGI-S score
3) PGI-C score
4) Proportion achieving threshold response < 30 eos/hpf in all of the 5 hpfs identified

5) Absolute change in endoscopic score measured by EoG-EREFS.
6) Absolute change in the total and each individual EoG-SQ eDiary symptom score
7) Absolute change in proportion of days with each individual symptom, multiple symptoms, and any (1 or more) symptom as measured by EoG-SQ eDiary
8) Absolute change in SODA, SF-12, PROMIS 29, PROMIS 25, and CGI-S scores.
9) CGI-C score
10) Induction of disease remission (percentage who achieve histological remission in the stomach as defined by peak eosinophil counts less than 30/hpf in 1, 2, 3 and 4 of the 5 hpfs identified for the primary endpoint).
11) Absolute change in EoG histology score (EoG HSS). For participants with esophageal involvement and/or duodenal eosinophilia involvement, absolute change in EoE HSS and EoD HSS, respectively.
12) Absolute and relative change in blood biomarkers, tissue biomarkers, and gastric transcripts
13) Absolute and relative change in blood eosinophil count, hemoglobin, hematocrit, total protein levels, albumin, IgE, iron storage panel (iron, ferritin, total iron-binding capacity)
14) The association of blood and tissue biomarkers and transcripts with drug responsiveness.
15) Absolute changes in esophageal and duodenal peak eosinophil levels.
16) Absolute changes in EoE-EREFS and EoD-EREFS in participants with esophageal and/or duodenal involvement, respectively.

Safety Endpoints

All AEs will be tracked and reported. Distributions of vital sign, urinalysis, blood chemistry and hematology measures will be evaluated.

Statistical Analyses

General Principles

Statistical analyses will be reported using summary tables, figures, and data listings. Continuous variables will be summarized with counts, means, standard deviations, medians, confidence intervals, minimums, and/or maximums. Categorical variables will be summarized by counts and by percentage of patients. Formal inferential statistical analyses techniques will be discussed in subsequent sections of this SAP. Statistical testing will be performed at the two-sided 5% level of significance.

All analyses will be performed using SAS Version 9.4 or higher on a PC platform. Table, listings, and figures will be presented in Word format. Upon completion, all SAS programs will be validated by an independent programmer. In addition, all program output will undergo a senior level statistical review. The validation process will be used to confirm that statistically valid methods have been implemented and that all data manipulations and calculations are accurate. Checks will be made to ensure accuracy, consistency with this plan, consistency within tables, and consistency between tables and corresponding data listings. Upon completion of validation and quality review procedures, all documentation will be collected and filed by the project statistician or designee.

Prior to statistical analysis, quality and distribution of the data will be examined. All individuals completing study visit 1 will be included in the analysis of the primary outcome including participants receiving steroids within the last 30 days before endoscopy. The patterns of data missingness will be evaluated by reviewing patient drop out reasons and comparing baseline patient characteristics. Missing outcome data will then be imputed using multiple imputation. Factors to be considered in the multiple imputation are the prior eosinophil counts, participant clinical characteristics and longitudinal secondary outcomes. If differences occur between the results of the ITT and PP analyses, we will explore the data to understand the pattern of missingness, using tipping point analysis to understand the impact of the missing data on inference, and consider additional imputation.

Major Protocol Violations

Major protocol violations will be identified by the clinical study team and provided to Biostatistics prior to database lock. A protocol deviation is any noncompliance with the clinical trial protocol or Good Clinical Practice (GCP). The noncompliance may be either on the part of the patient, the investigator, or the study site staff. All patients with major protocol violations will be listed by study center and patient numbers. A protocol deviation committee consisting of the project manager, study principal investigator, and lead statistician will review all cases prior to unblinded. Deviations will be determined without knowledge of randomization assignment prior to database lock.

Patient Enrollment and Disposition

Patient enrollment by site will be tabulated by treatment arm and overall. Patient disposition will be summarized by treatment arm and overall. This will include number of patients screened, number of patients who were screen failures with reason, number of patients who consented, and number of patients who were randomized. The summary will include the number and percentage of patients in each of the defined analysis populations. In addition, frequency counts and percentages of patients' reported reasons for ending the study will be summarized.

A listing will be presented to describe patient study arm, date of first and last dose, date of last visit or contact, and the reason for ending the study for each patient. Listings of inclusion/exclusion criteria responses will also be provided.

Description of Demographic and Baseline Characteristics

A summary of baseline characteristics (age, sex, race, ethnicity, atopy, historical EGIDs, years since diagnosis and concomitant steroid use) and baseline measures (gastric eosinophil count, histology scores, endoscopy scores, symptom scores, blood eos counts, and PRO scores) will be presented by treatment arm. Demographic/baseline characteristics and measures will be summarized within each treatment group using frequency and percentages for categorical variables and mean and standard deviation or median and interquartile range for continuous variables. The two treatment groups will be compared using Fisher's Exact Test for categorical variables and the nonparametric Wilcoxon Ranks Sum Test for the continuous variables. These summaries will include patients in the ITT population and PP population.

Efficacy Analyses

Primary Endpoint

The primary outcome measure will be the relative change in the mean of the eosinophil counts from the 5 worst microscopic fields in the gastric antrum and/or body. Relative change is defined as (week 12 – baseline)/baseline. Regression analyses will be performed with the primary outcome measure as the dependent variable and treatment group as the independent variable. Appropriate transformations, such as a log-transformation, may be used to satisfy the distributional assumptions prior to model fitting. If there are continued concerns about data distributions, equivalent nonparametric tests will be employed. To ensure that immunosuppression (e.g. corticosteroid, swallowed glucocorticosteroids for EoG (Entocort), or immunosuppressive agents) use does not influence results, we will also perform regression analyses with corticosteroid/immunosuppression therapy as a covariate (including considering binary (y/n), and dose (none, low, high).

Secondary Endpoints

For the secondary outcome measures, when normality assumption holds, we will test the association between outcomes and the arms of the study using linear mixed model; otherwise generalized linear mixed effects models will be used with the appropriate distribution function. Random study site effect will be tested; if significant, it will be included in the modeling; if insignificant, the models will be simplified to linear or generalized linear models. Appropriate transformations may be used to satisfy the normality assumptions of the model, if needed. Baseline corticosteroid/immunosuppressive therapy use will be included as a covariate in

the models. To evaluate factors associated with drug responsiveness, we will use models as described above but restrict the analyses to those in the medication arm of the study.

For the primary and secondary outcome measures, sensitivity analyses may be done by including the following covariates in the models described above: sex, age group, baseline eosinophil counts, missing last injection before EGD, and baseline disease severity. Given the sample size, covariates will be tested individually in the models.

Exploratory Endpoints

For the exploratory outcomes measured at baseline and week 12, when normality assumption holds, we will test the association between outcomes and the arms of the study using linear mixed model; otherwise generalized linear mixed effects models will be used with the appropriate distribution function. Random study site effect will be tested; if significant, it will be included in the modeling; if insignificant, the models will be simplified to linear or generalized linear models. Appropriate transformations may be used to satisfy the normality assumptions of the model, if needed. Baseline corticosteroid/immunosuppressive therapy use will be included as a covariate in the models.

Some exploratory outcomes may include measures during or at the end of OLE. For participants that continue into the open label phase of the study, we will evaluate change for two time ranges. We will examine the overall change across the study (baseline to OLE/EOT (Week 36)). Then, we will examine change during the open label phase (DBP/EOT (Week 12) to OLE/EOT (Week 36)). When evaluating these changes, analyses will be stratified by treatment status during the blinded phase of the study. Changes of the OLE endpoints will be tested against 0 using paired t tests or signed rank tests when lacking of normality. Mean change and two-sided 95% confidence intervals will be presented.

Safety outcomes

Adverse events will be summarized for each treatment group in the intent-to-treat population. The number and percentage with exact 95% confidence intervals of participants experiencing each adverse event category and the number of adverse events that occurred will be presented. Serious adverse events will be summarized using the same approach.

The following descriptive data displays and summaries will be provided for vital sign, urinalysis, blood chemistry and hematology measures. For continuous measures, scatter plots of each measure for each treatment group will be presented with the baseline measure on the x-axis and the week 12 measure on the y-axis. A 45-degree reference line will be included in each plot. For categorical measures, two-way frequency tables will be provided for each treatment group that include baseline result versus week 12 result. Abnormal (clinically significant) results will be reported. Any clinically relevant changes in physical exam data will be described for each treatment group.

Statistical Analysis Changes From the Protocol

Changed the analysis population for the key primary and secondary efficacy outcomes to the complete case ITT. The rationale was that a single participant in the placebo arm dropped out before the end of the double blind. The initial SAP stated that multiple imputation would be used if missing data occurred. However, multiple imputation considers a multitude of plausible values and accounts for this uncertainty in the modelling. While it is a statistically robust approach, graphing the observed data points would exclude the missing data thus providing a mismatch between the graphical data and the statistical models used. Further, there was striking agreement between the results using the imputed and complete case datasets.

Conventions

The precision of original measurements will be maintained in summaries, when possible. Means, medians and standard deviations will be presented with an increased level of precision; means and medians will be

presented to one more decimal place than the raw data, and the standard deviations will be presented to two more decimal places than the raw data.

Summaries of continuous variables that have some values recorded using approximate values (e.g., lab values < 0.001 or >200) will use imputed values. The approximate values will be imputed using the closest exact value for that measurement. For tables where rounding is required, rounding will be done to the nearest round-off unit. For example, if the round-off unit is the ones place (i.e., integers), values ≥ XX.5 will be rounded up to XX+1 while values < XX.5 will be rounded down to XX.

Percentages will be based on available data and denominators will generally exclude missing values. For frequency counts of categorical variables, categories whose counts are zero will be displayed for the sake of completeness. For example, if none of the patients discontinue due to "lost to follow-up," this reason will be included in the table with a count of 0. Categories with zero counts will not have zero percentages displayed.

Standard Calculations

EoG-EREFS

Will be calculated by location, by feature and as a total score. For fundus, body and antrum, the score will be the sum of the six features from each of the locations. For granularity, erosions, nodularity, erythema, friability and folds, the feature scores will be the sum of the feature of the three locations. The total score will be the sum of the six features of all three locations.

EoE-EREFS

Will be calculated by location, by features and as a total score. For proximal and distal location, the score will be the sum of all five features from each of the locations. For edema, rings, exudates, furrows, and stricture, the feature score will be the sum of the proximal and distal scores of the feature. The total score will be the sum of all features from proximal and distal locations.

EoD-EREFS

Will be calculated for duodenal bulb and duodenal, by features and as a total score. The score by location will be the sum of the seven features for the location, respectively. The feature score will be the sum of the feature from the two locations. The total score will be the sum of all features from both locations.

EoG-SQ

Total Symptom Score (TSS): TSS will be based on patient-reported severity scores of stomach pain, stomach cramping, nausea, bloating, early satiety, and loss of appetite, with each reported on a 0-10 scale. These individual scores are summed on each day, with maximum daily score of 60. The post-baseline TSS is then calculated by averaging daily sum scores over all days with eDiary data in a 7-day period. At least 4 of 7days must be entered in the EG-SQ eDiary to derive the post-baseline TSS. The baseline TSS will be calculated from the 14 days prior to V1 (week 0). At least 8 of 14 days must be entered in the EG-SQ eDiary to derive the baseline TSS.

Individual Symptom Scores (9 in total): For each individual score, the post-baseline value is calculated by averaging daily scores with eDiary data in a 7-day period. At least 4 of 7days must be entered in the EG-SQ eDiary to derive the post-baseline score. The baseline value will be calculated from the 14 days prior to V1 (week 0). At least 8 of 14 days must be entered in the EG-SQ eDiary to derive the baseline.

When multiple records are available for a patient on the same day, the last record will be used for analysis. The analysis visit window for symptom diary data is outline below:

	Analysis Visit Window
Baseline	Day -14 to Day -1
Week 1	Day 1 to Day 7
Week 2	Day 8 to Day 14

Week 3	Day 15 to Day 21
Week 4	Day 22 to Day 28
Week 5	Day 29 to Day 35
Week 6	Day 36 to Day 42
Week 7	Day 43 to Day 49
Week 8	Day 50 to Day 56
Week 9	Day 57 to Day 63
Week 10	Day 64 to Day 70
Week 11	Day 71 to Day 77
Week 12 Visit 5	Day 78 to Day 84 Day -7 to Day -1 since end of double blind.

Note:

- Diary study day = Diary date – 1st injection date in the study + 1 if diary date ≥ 1st injection date of study.
- Diary study day = Diary date – 1st injection date in the study if diary date < 1st injection date of study.
- Day 1 is the day of the first study drug injection, and Day -1 is the day before. There is no Day 0.
- The first dose of OLE is administered at Week 12, and any diary data taken after the first dose of OLE will be excluded from analysis.

SF-12

A SAS program will be used, which can be downloaded from https://labs.dgsom.ucla.edu/hays/pages/programs_utilities

PROMIS

For pediatric (age 8-17) and adult (18+) subjects, the scores will be calculated using PROMIS Pediatric Profile Instruments and PROMIS Adult Profile Instruments, respectively.

SODA

SODA scores will be generated as described in https://www.sciencedirect.com/science/article/pii/S0895435600003656?via%3Dihub

EoG HSS

Ratio of sum of all nonmissing features to the sum of maximum scores of all nonmissing features. There are 11 features, and the maximum score is 2 for each of the features. For the total score, all the 11 features will be used; for the inflammatory and structural sub-scores, the inflammatory and structural features will be used in the calculations, respectively. Inflammatory features include lamina propria eosinophil sheets, periglandular circumferential collars, eosinophils in surface epithelium, eosinophil glandulitis, eosinophil gland abscesses, eosinophils in muscularis mucosa, and acute inflammatory cells. Structural features include lamina propria fibroplasia, lamina propria smooth muscle hyperplasia, reactive epithelial changes, and surface erosion/ulceration.

EoE HSS

HSS will be the sum of grade and stage. Using the grade and stage scores for each of the eight features, grade and stage will be calculated as the ratio of sum of all nonmissing features to the sum of maximum scores of all nonmissing features. The maximum score is 3 for each of the features. HSS will be calculated for proximal and distal location, respectively. The maximum grade of the distal and proximal grade, and the maximum stage of the distal and proximal stage will then be summed to generate the maximum HSS.

EoD HSS

Ratio of sum of all nonmissing features to the sum of maximum scores of all nonmissing features. There are 13 features, and the maximum score is 2 for each of the features.

Study Design/Visits