A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of AK002 in Patients with Moderately to Severely Active Eosinophilic Gastritis and/or Eosinophilic Duodenitis (formerly referred to as Eosinophilic Gastroenteritis) Who Have an Inadequate Response with, Lost Response to, or Were Intolerant to Standard Therapies

NCT Number: NCT04322604
Document Date: 05 Nov 2021

Statistical Analysis Plan for Protocol AK002-016

Protocol Title	A Phase 3, Multicenter, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of AK002 in Patients with Moderately to Severely Active Eosinophilic Gastritis and/or Eosinophilic Duodenitis (formerly referred to as Eosinophilic Gastroenteritis) Who Have an Inadequate Response with, Lost Response to, or Were Intolerant to Standard Therapies
Protocol Number	AK002-016
Protocol Version	Amendment 4
Protocol Dates	**Amendment 4 26 October 2021** Amendment 3 20 April 2021 Amendment 2 07 August 2020 Amendment 1 18 March 2020 Original 30 September 2019
Study Drug	AK002 (lirentelimab)
Study Phase	3
IND Number	135158
Sponsor	Allakos Inc., 975 Island Drive, Suite 201, Redwood City, CA 94065 USA
SAP Version	1
SAP Date	**05 November 2021**

Signature Page

Prepared by:

PPD

06-Nov-2021 | 19:29 PDT

PPD PhD
PPD Principal Biostatistician
Pharma Data Associates, LLC

Date

Reviewed and Approved by:

PPD

05-Nov-2021 | 15:23 PDT

Director, Biometrics
Allakos, Inc.

Date

PPD

05-Nov-2021 | 15:31 PDT

PPD Biometrics
Allakos, Inc.

Date

PPD

06-Nov-2021 | 09:07 PDT

Allakos, Inc.

Date

Table of Contents

List of In-Text Tables

List of Abbreviations

ADA	Anti-drug antibody
ADaM	Analysis Data Model T
AE	Adverse event
ANCOVA	Analysis of covariance
ATC	Anatomical Therapeutic Chemical (Classification System)
BLOQ	Below limit of quantification
BMI	Body mass index
CBC	Complete blood count
CDISC	Clinical Data Interchange Standards Consortium
CMH	Cochran-Mantel-Haenszel
CRF	Case Report Form
CRO	Contract Research Organization
CSR	Clinical study report
DM	Data management
eCDF	empirical cumulative distribution function
ECG	Electrocardiogram
eCRF	electronic Case Report Form
EDC	Electronic Data Capture (system)
EG	Eosinophilic gastritis
EGD	Esophago-gastro-duodenoscopy
EGE	Eosinophilic gastroenteritis
EoD	Eosinophilic duodenitis
EoE	Eosinophilic esophagitis
ET	Early Termination
FDA	Food and Drug Administration
FSH	Follicle-Stimulating Hormone
hpf	High power field
ICE	Intercurrent events
ICH	International Conference on Harmonization
ICF	Informed Consent Form
IRT	Interactive Response Technology
ITT	Intent-to-treat (population)

IV	Intravenous
LLN	Lower limit of normal
LLOQ	Lower limit of quantification
LSM	Least squares mean
MAR	Missing at random
MCAR	Missing completely at random
MCMC	Markov Chain Monte Carlo (method)
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple imputation(s)
MITT	Modified intent-to-treat (population)
MMRM	Mixed Model for Repeated Measures
MNAR	Missing not at Random
p-value	Probability value
PD	Pharmacodynamics
PE	Physical examination
CCI	[redacted]
CCI	[redacted]
PK	Pharmacokinetic(s)
PP	Per protocol (population)
PPI	Proton pump inhibitor(s)
PRO	Patient-reported outcome
PT	Preferred term
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SAS	Statistical analysis system
SD	Standard deviation
SE	Standard error(s)
SDTM	Study Data Tabulation Model
CCI	[redacted]
SOC	System organ class
TEAE	Treatment-emergent adverse event(s)
TEAESI	Treatment-emergent adverse event(s) of special interest
TESAE	Treatment-emergent serious adverse event(s)
TLF	Tables, listings, and figures

TSS	Total Symptom Score
ULN	Upper limit of normal
WHO	World Health Organization
WHODD	World Health Organization Drug Dictionary

Revision History

Version Date	Version Number	Description
05 Nov 2021	**1**	Initial document

1. Introduction

This statistical analysis plan (SAP) describes the rules and conventions to be used in the analysis and presentation of efficacy and safety of AK002 in subjects with moderate-to-severe eosinophilic gastritis (EG) and/or eosinophilic duodenitis (EoD) who have an inadequate response with, lost response to, or were intolerant to standard therapies, as planned for the clinical protocol.

The SAP describes the data and variables to be summarized or analyzed, including specifications of the analytical methods to be performed. This SAP supersedes the statistical analysis methods described in the clinical protocol except for the standard pharmacokinetic (PK) data analyses. Significant deviations/changes from the planned analyses described in this SAP will be identified, with justification, in the appropriate section of the clinical study report (CSR). The SAP is based on Clinical Study Protocol AK002-016 Amendment 4, dated 26 October 2021, and the associated electronic case report forms (eCRF).

2. Study Objectives

2.1 Primary Objective – Primary Endpoints

The primary objectives of the study are to evaluate the efficacy and safety of 6 doses of AK002 in subjects with moderate-to-severe EG and/or EoD when compared with placebo.

Efficacy will be evaluated by 2 co-primary endpoints:

1) First co-primary endpoint – Proportion of Tissue Eosinophil Responders at Week 24:
   A responder is a subject achieving the following eosinophil counts:
   - EG: Mean eosinophil count ≤4 cells/hpf in 5 highest gastric hpf
   - EoD: Mean eosinophil count ≤15 cells/hpf in 3 highest duodenal hpf
   - EG + EoD: Mean eosinophil count ≤4 cells/hpf in 5 highest gastric hpf and mean eosinophil count ≤15 cells/hpf in 3 highest duodenal hpf
2) Second co-primary endpoint: Change in TSS from baseline to Weeks 23–24 as measured by the patient reported outcome (PRO) questionnaire.

   The PRO Total Symptom Score (TSS) comprises the following 6 symptoms:
   - Abdominal pain intensity
   - Nausea intensity
   - Fullness before finishing a meal intensity

- Loss of appetite intensity
- Bloating intensity
- Abdominal cramping intensity

2.2 Secondary Objectives – Secondary Endpoints

The secondary objectives are to characterize further the efficacy of AK002 in subjects with EG and/or EoD as measured by:

- Change in tissue eosinophils from baseline to Week 24
- Proportion of subjects achieving mean eosinophil count ≤1 cell/hpf in 5 highest gastric hpf and/or mean eosinophil count ≤1 cell/hpf in 3 highest duodenal hpf at Week 24
- Proportion of treatment responders at Week 24. Responder is defined as >30% improvement in TSS symptom score and mean eosinophil count ≤4 cells/hpf in 5 highest gastric hpf and/or mean eosinophil count ≤15 cells/hpf in 3 highest duodenal hpf
- Proportion of subjects who achieve ≥50% reduction in TSS from baseline to Weeks 23–24
- Proportion of subjects who achieve ≥70% reduction in TSS from baseline to Weeks 23–24
- Change in weekly TSS over time

2.3 Exploratory Objectives – Exploratory Endpoints

The exploratory objectives are to evaluate the effect of AK002 by comparing AK002 to placebo treatment for the following parameters:

- Change from baseline in CCI [redacted] over time
- Change from baseline in CCI [redacted] over time
- For subjects with CCI [redacted] Proportion of subjects achieving a CCI [redacted] at Week 24
- Change from baseline in CCI [redacted] over time

2.4 Safety Objectives – Safety Endpoints

The safety objective of the study is to evaluate the study drug safety profile using the following safety endpoints:

- Treatment emergent adverse events (TEAE) including severity, relationship to study treatment, action taken, and outcome, serious adverse events (SAE) and adverse events (AE) leading to study drug withdrawal
- Anti-drug (AK002) antibody (ADA)
- Blood chemistry
- Hematology
- Urinalysis
- Physical examination
- Changes in vital signs
- Changes in concomitant medication use due to AE

3. Study Design

3.1 General Description

This is a Phase 3, multicenter, randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of AK002 in subjects with moderate-to-severe EG and/or EoD who have an inadequate response, lost response, or were intolerant to standard therapies. Subjects enrolled in the study will receive 6 infusions of placebo or AK002 administered every 4 weeks and will be followed for 12 weeks after the last dose unless subjects elect to enter the optional long-term extension study.

Subjects who have signed the informed consent form (ICF) will be screened within 35 days prior to Study Day 1.

The study will be carried out as follows:

- An 18 to 45-day screening period with baseline evaluations, including active symptoms of disease (gathered by PRO questionnaire) and EGD with biopsy.
- Prior EGD may be used for eligibility as long as it was collected within 30 days of the first screening visit for the AK002-016 study and was performed and centrally evaluated for the AK002-016 study.
- If subjects meet histology and symptom eligibility criteria, they will be randomized after being stratified by the highest weekly TSS of disease activity recorded during the screening period (<28 or ≥28 strata) and whether the subject has EG with or without EoD (EG ± EoD) or EoD without EG. The interactive response technology (IRT) will randomly assign subjects 1:1 to receive AK002 or placebo.

- Pre-study medications and dietary restrictions will remain unchanged throughout the study. Systemic or topical steroids above 10 mg prednisone will not be allowed except due to unforeseen medical circumstances where it is deemed to be medically necessary to treat an unrelated medical condition or when given as a premedication prior to infusion or to treat an IRR that occurs during infusion (see Section 6.3 and Protocol Section 8.1).
- Eligible subjects will receive the first dose of placebo or AK002 (1 mg/kg) on Day 1, with premedication of 60 mg oral prednisone 12–24 hours prior to the start of the infusion.
- If the study drug is well tolerated (no stopping rules being met), subjects will receive additional doses of placebo or AK002 (3 mg/kg) on Days 29, 57, 85, 113, and 141. With the exception of Day 1, steroid premedication will only be allowed with the written approval of the Medical Monitor.
- Subjects will remain at the site for at least 1 hour of observation following the end of the infusion.
- During the Treatment period, subjects will return to the clinic for study visits as described in the Schedule of Events (Table 1).
- Subjects will continue to complete daily PRO questionnaires throughout the study and the follow-up period.
- A repeat EGD with biopsy will be performed on Day 169 (±3) or 28 (±3) days after last dose of study drug if subject is terminated early.
- Follow-up will occur for 84 (±3) days after the last dose unless subjects decide to enter a long-term extension study. Follow-up visits for subjects opting not to enter the extension study will occur on Days 176 (±3), 197 (±3), and 225 (±3).
- Subjects who satisfactorily complete through Day 176 of AK002-016 have the option to receive AK002 in a separate open-label extension study if all eligibility criteria for the extension study are satisfied. Subjects who enroll in the AK002-016X extension study prior to Day 197 occurring will not complete the Day 197 or Day 225 procedures under Protocol AK002-016.
- Total study duration is approximately 35–37 weeks. For subjects entering the AK002-016X extension study, the total study duration is approximately 28–37 weeks.

The overall schedule of procedures and assessments are presented in Table 1.

Table 1 Schedule of Assessments

	Screening (18–45 days)		**Treatment Period** (20 weeks)								**Follow-Up Period**[30] (12 weeks)			
	Baseline[1]	**Prior to Day 1**	**Dose 1 Day 1**	**Day 8** (±2 days)	**Day 15** (±2 days)	**Dose 2 Day 29** (±3 days)	**Dose 3 Day 57** (±3 days)	**Dose 4 Day 85** (±3 days)	**Dose 5 Day 113** (±3 days)	**Dose 6 Day 141** (±3 days)	**Day 169** (±3 days) or 28 days after last dose if ET	**Day 176** (±3 days) or 35 days after last dose if ET[29]	**Day 197** (±3 days) or 56 days after last dose if ET[29]	**Day 225** (±3 days) or 84 days after last dose if ET[29]
Informed consent	X													
Demographics	X		X											
Medical History	X		X											
Prior/Concomitant Medications	X		X			X	X	X	X	X	X	X	X	X
Body weight and height[2]	X		X			X	X	X	X	X		X	X	X
Vital Signs[3]	X		X			X	X	X	X	X		X	X	X
10 or 12-lead ECG[4]	X													
Complete Physical Exam[5]	X													
Symptom-Directed Physical Exam[6]			X			X	X	X	X	X		X	X	X
Baseline Diet Assessment[7]	X													
Baseline Diet Compliance[8]			X			X	X	X	X	X		X	X	X
Stool for Ova and Parasite[9]	X													
ePRO Activation and Training[10]	X													
ePRO Questionnaire (may include Additional Questions)[11]	<----------Complete electronically one time daily ---------->													
CCI [redacted]	X		X			X	X	X	X	X		X	X	
CCI [redacted]	<----------Complete electronically on Screening Day 19, Study Day 7, and Study Day 28 ---------->													
CCI [redacted]	<----------Complete electronically on Study Day 7 and Study Day 28---------->													
EGD with Biopsy[12,14]	X										X			
Blood for Screening Chemistry (includes hCG and FSH)[12,15]	X													

Table 1 Schedule of Assessments cont.

	Screening (18–45 days)		**Treatment Period** (20 weeks)								**Follow-Up Period**[30] (12 weeks)			
	Baseline[1]	**Prior to Day 1**	**Dose 1 Day 1**	**Day 8** (±2 days)	**Day 15** (±2 days)	**Dose 2 Day 29** (±3 days)	**Dose 3 Day 57** (±3 days)	**Dose 4 Day 85** (±3 days)	**Dose 5 Day 113** (±3 days)	**Dose 6 Day 141** (±3 days)	**Day 169** (±3 days) or 28 days after last dose	**Day 176** (±3 days) or 35 days after last dose if ET[29]	**Day 197** (±3 days) or 56 days after last dose if ET[29]	**Day 225** (±3 days) or 84 days after last dose if ET[29]
Blood for Serology and *Strongyloides stercoralis* [12,16]	X													
Blood for Total Serum IgE[12,17]	X											X		
Previous Treatments Review	X													
Eligibility Assessment	X	X	X											
Access IRT: Stratification and Randomization[18]			X											
Access IRT: IP Kit Assignment			X			X	X	X	X	X				
Premedication: Prednisone[19]		X												
Study Drug Administration[20]			X			X	X	X	X	X				
Blood for CBC with differential[12,21]	X		X	X	X	X	X	X	X	X	X	X	X	X
Blood for Chemistry[12,22]			X	X	X	X	X	X	X	X	X	X	X	X
Blood for PK[12,23]	X			X	X	X	X	X	X	X	X	X	X	X
Blood for ADA[12,24]	X			X	X	X	X	X	X	X	X	X	X	X
Urine for Urinalysis[12,25]	X		X			X	X	X	X	X		X	X	
Dipstick Pregnancy Test[26]			X			X	X	X	X	X		X	X	
Non-serious Adverse Events[27]			X			X	X	X	X	X	X	X	X	X
Serious Adverse Events[28]			X			X	X	X	X	X	X	X	X	X
Begin AK002-016X extension study at least 1 day after Day 176 Visit (if applicable)[31]												X[33]	X[33]	X[33]
Blood for histamine/tryptase[32]			X			X	X	X	X	X				

Table 1 Notes

ADA: Anti-AK002 antibody
CBC: Complete blood count
ECG: Electrocardiogram
ePRO: electronic Patient Reported Outcome
ET: Early Termination
FSH: Follicle-stimulating hormone
hCG: Human Chorionic Gonadotropin
IP: Investigational Product
IRT: Interactive Response Technology
CCI [redacted]
PK: Pharmacokinetics

1) Baseline screening visit can occur over several days within the screening period. Day 1 can begin as soon as eligibility criteria are met and at least 4 diary entries have been completed during the third week of screening.

2) At screening, height (in cm) and weight (in kg) will be recorded. Body weight will also be measured on Days 1, 29, 57, 85, 113, 141 and on follow-up Days 176, 197, and 225 or 28, 56, and 84 days after last dose if ET. Current body weight or body weight from 1 day prior will be used to calculate the amount of AK002/placebo to be mixed with NaCl for the appropriate dose to be administered on each infusion day.

3) Vital signs will be measured at screening, Days 176, 197, and 225 or 28, 56, and 84 days after last dose if ET and on all dosing days: within 30 minutes predose, 15 minutes (±5 minutes) after the start of study drug infusion, within 15 minutes following the end of infusion and just prior to discharge. Additional vital signs measurements may be collected at the Investigator's discretion if an IRR occurs. Vital signs including systolic and diastolic blood pressure, pulse, body temperature, and respiratory rate will be measured after the patient has been at rest for ≥5 minutes and before any blood draws have been obtained (unless collected for an IRR).

4) A 10-lead or 12-lead ECG will be obtained at screening before any blood is drawn and after the patient has been in the appropriate position for ≥5 minutes.

5) A complete physical examination will be performed by either the Investigator or designee and include the following body system or organ assessments: skin; head, eyes, ears, nose and throat; thyroid; lungs; cardiovascular; abdomen; extremities; lymph nodes; and a brief neurological examination.

6) A symptom-directed PE (including assessment of possible infusion site reactions) will be performed by the Investigator or designee, as needed if any symptoms are reported.

7) A baseline diet assessment (Protocol Appendix 3) will be performed using standardized questions. Eating patterns, food avoidance behaviors, and confirmed allergies will be captured.

8) A baseline diet compliance check will be performed at every study visit, except as noted above, and any variances from the baseline diet documented. Patients should maintain the baseline diet throughout the study.

9) Fecal collection kits for Ova and Parasite test will be provided to patients at screening. Collection kits should be returned to the clinical site within 1 day of collection. Negative results must be available prior to randomization.

10) Activate PRO questionnaire and provide patient with unique username and password. PRO questionnaire should be activated for all patients on screening Day 1. Patients with concomitant history of asthma or atopic dermatitis will receive an extra question, about each, as appropriate. All patients will receive dysphagia question during screening and will continue to receive the question after enrollment if weekly average dysphagia is ≥3 during any 2 weeks of screening.

11) PRO should be completed around the same time each day. Prior to enrollment, the PRO weekly averages of abdominal pain, nausea, and diarrhea over the screening period will be calculated and used to assess eligibility. Weekly average Total Symptom Score (TSS6) will be calculated for eligibility and stratification. PRO weekly average for dysphagia will also be assessed prior to enrollment to determine if the patient should continue receiving the 1-question daily dysphagia question.

Table 1 Notes cont.

12) Specimen processed by central laboratory. See central laboratory manual for collection and processing details.

13) To be completed electronically by patient, in clinic, prior to any blood draw, physical exam, or vital sign measurements.

14) See Protocol Appendix 5 for biopsy assessments. The post-treatment endoscopy (EGD) and biopsy assessments will be performed on Day 169 (±3) or 28 (±3) days after last dose of study drug if ET. If the Day 141 infusion occurs out of window, the EGD should occur 28 (±3) days from the date of Day 141. Post-treatment EGD biopsy results will be blinded to the site.

15) Blood for baseline Chemistry, including hCG and FSH, will be collected during the screening period. Only patients of childbearing potential and post-menopausal women are required to have hCG and FSH testing completed.

16) Blood for Serology testing will be collected during screening and will include HBsAg, hepatitis C antibody, anti-HBc, and HIV, as well as *Strongyloides stercoralis*.

17) Blood samples for Total Serum IgE will be collected during screening and on Day 176 or 35 (±3) days after last dose of study drug, if ET.

18) Stratification based on TSS of <28 or ≥28 and whether patient is EoD-only will occur. Randomization and stratification will be conducted through the IRT system.

19) Premedication with 60 mg oral prednisone or approved alternative is required 12–24 hours prior to first infusion. Premedication may be administered prior to subsequent infusions at the Investigator's discretion, but only with written Medical Monitor approval.

20) Study drug will be administered as a single peripheral IV infusion over at least 4 hours for Dose 1 and at least 3 hours for Dose 2. Doses 3–6 should occur over at least 2 hours, depending on prior infusion tolerability. Please refer to the Pharmacy Manual.

21) Blood for CBC with differential, including absolute blood eosinophil count, will be obtained just prior to each infusion, 1 hour (±15 minutes) after the end of each infusion as well as during the screening period, and on Days (±3) 8, 15, 169, 176, 197, and 225 or 35, 56, and 84 (±3) days after last dose if ET. All differential blood counts from Day 1 (post-dose) through the end of the patient's participation will be blinded to the Sponsor and the site. An unscheduled CBC may be collected at the request of the Safety Monitor.

22) Blood for chemistry will be obtained predose on dosing Days 1, 29, 57, 85, 113, and 141, as well as on Days 8, 15, 169, 176, 197, and 225 or 28 (±3), 35 (±3), 56 (±3), and 84 (±3) days after last dose of study drug if ET.

23) Blood for PK will be obtained predose on dosing Days 29, 57, 85, 113, and 141, as well as during screening and on Days 8, 15, 169, 176, 197, and 225 or 28 (±3), 35 (±3), 56 (±3), and 84 (±3) days after last dose of study drug if ET.

24) Blood for ADA will be collected predose on dosing Days 29, 57, 85, 113, and 141, as well as during screening and on Days 8, 15, 169, 176, 197, and 225 or 28 (±3), 35 (±3), 56 (±3), and 84 (±3) days after last dose of study drug if ET. ADA sample will also be collected any time an immunogenicity-related AE occurs.

25) Urine for standard urinalysis will be obtained predose on dosing Days 1, 29, 57, 85, 113, and 141 as well as during the screening period and on Day 176 and Day 197 or 35 (±3) days and 56 (±3) days after last dose of study drug if ET.

26) A small amount of urine collected for urinalysis will be used for dipstick pregnancy test on all infusion days and on Day 176 and 197 or 35 and 56 (±3) days after last dose of study drug, for all patients of childbearing potential. Test kits will be supplied by the central laboratory. Tests will be completed on site and evaluated prior to each infusion.

Table 1 Notes cont.

27) The capture of non-serious AE and adverse events of special interest will begin after the first dose of study drug has occurred.

28) The reporting of SAE occurring after signing Informed consent and prior to the first infusion will be limited to those that relate to screening procedures. The capture of all SAE and AE that are not related to screening procedures will begin at the time of first infusion of study drug.

29) Patients who sign the informed consent for the AK002-016X extension study will complete the procedures for Day 176 and may begin open-label dosing in the extension study, if eligible, one day after the completion of Day 176 visit. In this case, patients will not complete the Day 197 or Day 225 procedures under Protocol AK002-016. Open-label dosing and follow-up will occur under the AK002-016X extension study. Patients who choose to enter the AK002-016X extension study after Day 197 but before Day 225 will not complete Day 225 procedures.

30) The ET visits should be conducted 28 (±3), 35 (±3), 56 (±3), and 84 (±3) days after the last dose of study drug or prior to this, if necessary, to ensure compliance with the visit. If only 1 ET visit is possible, EGD and end-of-study blood work may occur on the same day (28 days after last dose and 35 days after last dose visit procedures). If the end-of-study visit occurs more than 35 days after the last dose of study drug, then perform the visit as soon as possible. The procedures listed under the 28-day and 35-day post-study drug visit will be conducted unless otherwise directed by the Medical Monitor.

31) The final hematology assessment for AK002-016 ***must*** be collected ***prior to*** the subject taking prednisone premedication for AK002-016X. Therefore, dosing in AK002-016X must take place at least 1 day after completion of the Day 176 visit for AK002-016. Subjects may start AK002-016X dosing 1 day after completion of the Day 176 Visit.

32) Blood samples for histamine and tryptase to be collected within 1–2 hours of start of symptoms if there is any suspicion of anaphylaxis.

33) For patients who do not start the extension study before Day 197 or Day 225, visit procedures will occur as per the AK002-016 study.

3.2 Study Treatment

3.2.1 Treatment, Dose, and Mode of Administration

Subjects will be randomly assigned through the IRT system to an active dose group of 1 mg/kg AK002 followed by 5 doses of 3 mg/kg or placebo. The exact dose will be calculated prior to each infusion and based on current subject weight. Study drug will be administered as a single peripheral IV infusion using an infusion pump on Days 1, 29 (±3), 57 (±3), 85 (±3), 113 (±3), and 141 (±3).

3.2.2 Duration of Study

- Screening phase: 18–45 days prior to study drug administration
- Treatment phase: 20 weeks
- Follow-up period: 84 days (±3) following the last dose of study drug

3.2.3 Methods of Assigning Subjects to Treatment Group

Approximately 160 subjects with symptomatic EG and/or EoD will be stratified by baseline TSS and disease status (EG ± EoD, EoD). Subject will be randomized 1:1 to receive 1 of 2 dose regimens using stratified randomized blocks in a double-blind manner.

- AK002 at 1 mg/kg for the first dose followed by 3 mg/kg administered every 4 weeks for 5 subsequent doses.
- 6 doses of placebo

It is anticipated that at least 90 subjects with EG ± EoD will be enrolled, and up to 70 subjects with EoD without EG will be enrolled.

3.3 Blinding

This is a double-blind study. The identity of active and placebo treatments will not be known to Investigators, Sponsor (including safety monitor), research staff (including pharmacy), subjects, or the study monitor.

3.4 Hypotheses

The hypothesis to be tested in the study is that AK002 is different from placebo with regard to the 2 co-primary efficacy endpoints of the proportion of subjects with tissue eosinophil response at Week 24 and the mean reduction from baseline in TSS at Weeks 23–24.

3.5 Determination of Sample Size

A total of approximately 160 subjects will be enrolled.

First Co-Primary Endpoint: A sample size of 80 subjects per treatment group will have >99% power to demonstrate a greater proportion of tissue responders at Week 24 in AK002 subjects when compared to placebo subjects, assuming the proportions of responders are 0.6 and 0.1 in AK002 and placebo groups, respectively.

Second Co-Primary Endpoint: A sample size of 80 subjects per group will provide 96% power to detect a statistically significant difference of 7.4 points between AK002 and placebo in the mean reduction from baseline in TSS at Weeks 23–24, assuming a common standard deviation of 12.5 points and a population mean baseline TSS of 30 points.

4. Definitions

4.1 Terminology and Definitions

Table 2 Terminology and Definitions

Terminology	Definition
"And/Or"	Condition A "and/or" B is True when • A is True and B is not evaluable (not applicable) • B is True and A is not evaluable (not applicable) • A is True and B is True when both A and B are evaluable
Baseline	Baseline for non-daily assessment (e.g., laboratory tests and CCI) is defined as the non-missing value collected most recent to and before the time of the very first dose of study drug. This includes lab test collected on Day 1 Predose as an example. Baseline Total Symptom Score (TSS) will be the average of the weekly TSS collected in the last 2 weeks prior to the first dose.
Completer for the Study	Subjects who complete at least through the Day 176 visit if continuing to the AK002-016X study or Day 225 visit if not continuing to the AK002-016X study.
Concomitant Medication	Medication collected on the Prior/Concomitant Medication CRF, with end date on/after Study Day 1. Note a Prior Medication may also be a Concomitant Medication if the start date is prior to Study Day 1 and end date is on/after Study Day 1.
Enrolled	Subject who is randomized to a treatment group.
Newly Initiated Medication	Newly initiated medication refers to any medication with a start date ≥ Study Day 1
Patient-Reported Outcome	The patient-reported outcome (PRO) questionnaire evaluates 8 different symptoms with 10 daily questions for intensity and frequency: • Abdominal pain intensity • Nausea intensity • Vomiting intensity • Vomiting frequency

Table 2 Terminology and Definitions cont.

Terminology	Definition
Patient-Reported Outcome cont.	• Diarrhea intensity • Diarrhea frequency • Fullness before finishing a meal (Early satiety) intensity • Loss of appetite intensity • Bloating intensity • Abdominal cramping intensity Each intensity evaluation is scored on a scale of 0=none to 10=worst possible.
Prior Medication	Medication collected on the Prior/Concomitant Medication CRF, with start date prior to Study Day 1.
PRO Total Symptom Score	Total symptom score (TSS) is the sum of 6 (abdominal pain, abdominal cramping, bloating, nausea, early satiety, and loss of appetite) weekly average symptom intensity scores.
Study Day	Study Day 1 is defined as the date on which a subject took the first dose of study drug. Other study days are defined relative to Study Day 1. For visits prior to the first dose of study drug, Study Day is calculated as Visit Date – Day 1 Date. For visits after the first dose, Study Day is calculated as Visit Date – Day 1 Date +1.
Study Drug	AK002 or placebo administered by IV infusion.
Study Week	Study Week for PRO analysis is defined as 7 days a week starting from the day of first dose (Day 1).
Tissue Eosinophil Responder in Disease Status (2 levels; EG ± EoD, EoD)	For EG with/without EoD subjects: Mean eosinophil count ≤4 cells/hpf in 5 highest gastric hpf and/or mean eosinophil count ≤15 cells/hpf in 3 highest duodenal hpf. For EoD subjects: Mean eosinophil count ≤15 cells/hpf in 3 highest duodenal hpf.
Treatment-emergent	Adverse events reported in the clinical database with a date of onset on or after the start date of the first dose of the study drug.
Treatment Responder	Subject with >30% improvement in TSS from baseline and mean eosinophil count ≤4 cells/hpf in 5 highest gastric hpf and/or mean eosinophil count ≤15 cells/hpf in 3 highest duodenal hpf.

4.2 Target of Estimation

The estimand (target of estimation) for Protocol AK002-016 is:

In subjects with EG and/or EoD, what is between group (AK002 vs. Placebo) difference in the proportion of tissue eosinophil responders at Week 24, and group difference in TSS change from baseline to Weeks 23–24 as measured by the patient-reported outcome (PRO) questionnaire.

The sections below describe the attributes of the estimand consistent with the ICH E9 (R1) Addendum (FDA, 2021).

4.2.1 Population Targeted by the Scientific Question

The population targeted by the scientific question is defined by the inclusion and exclusion criteria as part of the study protocol. Subjects must have a histological diagnosis of EG and/or EoD, as well as moderately to severely active symptomatic disease.

A key aspect of eligibility is that subjects must complete at least 4 daily PRO questionnaires per week for a minimum of 3 weeks during the screening period. In addition, subjects must have a weekly average score of abdominal pain, nausea, or diarrhea ≥3 on the PRO questionnaire (score from 0–10) for at least 2 weeks of screening and a weekly average TSS of ≥10 for at least 2 weeks of screening.

4.2.2 Variables of Interest (or Endpoint) to be Obtained for Each Subject that is Required to Address the Scientific Question

The co-primary endpoints to be obtained for each subject in this study to address the scientific question are tissue eosinophil responders (as defined in Table 2) at Week 24 and change in TSS from baseline to Weeks 23–24 as measured by the PRO questionnaire.

4.2.3 Treatment

AK002 or placebo administered to subjects on Days 1, 29, 57, 85, 113, and 141.

4.2.4 Intercurrent Events

The events below are considered intercurrent events (ICE) confounding with the efficacy outcomes.

- Premature discontinuation from the study
- Use of prohibited/restricted medication

Further clarification and handling of ICE including prohibited/rescue medications is detailed in Appendix 2.

4.2.5 Strategy for Handling Intercurrent Events

For the analysis of the study product estimand, tissue eosinophil values and TSS will be counted as non-responders for binary variables and set to missing for continuous outcomes from the point when an ICE occurs. An appropriate method for handling missing data through statistical modeling (e.g., multiple imputation [MI]) will be used (Sections 5.4.1, 5.4.2, and 6.11). The estimand will provide an answer to the question that is crucial to individual subjects:

"If I take this study drug as part of my treatment regimen, without adding any further medications that may impact the underlying disease or exit the study prematurely, what improvements in histology and PRO symptoms might be anticipated after 24 weeks?"

4.2.6 Summary Measure of the Estimand

- Percent (and 95% CI) of subjects having tissue eosinophil response at Week 24 in the AK002 and Placebo treatment groups and the absolute difference (and 95% CI) in the percent response between treatments.
- Least squares mean (LSM) (and standard error [SE]) of change from baseline to the average of Weeks 23–24 TSS and the between treatment difference in the AK002 and placebo treatment groups LSM.

5. Statistical Methods

5.1 General Methodology

All statistical analyses will be conducted using SAS v9.4 or later version on the Microsoft Windows Operating System.

All CRF data (raw data) will be converted into SDTM (Study Data Tabulation Model) datasets, which will be used to create ADaM (Analysis Data Model) data sets. The creation of the SDTM and ADaM data sets will follow the CDISC (Clinical Data Interchange Standards Consortium) standards and the FDA Study Data Technical Conformance Guide. All analysis tables and listings will be created from the ADaM data sets.

Continuous data will be summarized using "n" (number of subjects with non-missing observations), mean, median, standard deviation (SD), minimum value, and maximum value. Categorical data will be summarized using the frequency count and percentage (n, %) of subjects in each category. Number of subjects with non-missing values or number of subjects with missing values (e.g., Not Done) will be presented, where appropriate. Subjects with missing values will not contribute to the denominator for percentage calculations, unless specified otherwise. Counts of 0 in any category will be presented without percentage. All summaries will be presented for individual treatment groups. In addition, for summary of disposition and subject baseline characteristics, the presentation will include both treatment groups combined.

The precision rules for the presentation of summary statistics will be:

- Sample size (n, N) and number of missing responses (if displayed): Integer
- Mean, confidence interval, and median: Same number of decimal places as reported/collected
- Standard deviation: Same number of decimal places as reported/collected
- Percentiles, minimum, maximum: Same number of decimal places as reported/collected
- Odds Ratio: 2 decimal places
- Percentage: 1 decimal place generally, or 2 decimal places for <0.1%, or no decimal places for 0% and ≥100%
- P-value: 4 decimal places
- WBC: 2 decimal places as 0.01×10^9/L
- Height/Weight/BMI: 1 decimal place

The data summaries will be accompanied by individual subject data listings. All data available from questionnaires, eCRF, and external transfer (labs) will be listed and will include relevant subject information, e.g., treatment group and study day. The listings will be sorted in the order of treatment group, subject ID, assessment name and date/time.

Dates will be presented in the ISO-8601 format YYYY-MM-DD. Times will be displayed in 24-hour clock format. Numbering for tables, figures and listings will follow ICH E3 Guideline (ICH, 1996).

Alternative methods of analysis of the data may be considered prior to database lock should some of the assumptions underlying the proposed analyses not be met. Reason for departure from the planned methods will be documented as an amendment to the SAP or in the CSR.

5.2 Visit Window and Unscheduled Assessments

Data collected for study assessments provide information on the status of the subject at a given time point. These may provide biased results if the assessment is performed early or late. Therefore, assessments will be slotted into analysis windows to allow summaries to be performed for subjects with similar study drug exposure. The analysis window is constructed by the medians of 2 target study days of the adjacent planned visits.

Table 3 Analysis Window Rules for CCI [redacted]

Visit Description	Window
Baseline	On or Prior to Day 1 Predose
Day 1	Predose on infusion Day 1
Day 29	Day 1 Postdose to Day 42
Day 57	Day 43 to Day 70
Day 85	Day 71 to Day 98
Day 113	Day 99 to Day 126
Day 141	Day 127 to Day 158
Day 176	Day 159 to Day 186
Day 197	Day 187 to End of Study

Table 4 Analysis Window Rules for Vital Sign and Laboratory Tests

Visit Description	Window
Baseline	On or Prior to Day 1 Predose
Day 1	Day 1 Postdose
Day 15	Day 2 to Day 22
Day 29	Day 23 to Day 42
Day 57	Day 43 to Day 70
Day 85	Day 71 to Day 98
Day 113	Day 99 to Day 126
Day 141	Day 127 to Day 158
Day 176	Day 159 to 186
Day 197	Day 187 to End of Study

Table 5 Analysis Window Rules for PK Concentration and ADA

Visit Description	Window
Baseline	On or Prior to Day 1 Predose
Day 8	Day 1 Postdose to Day 11
Day 15	Day 12 to Day 22
Day 29	Day 23 to Day 42
Day 57	Day 43 to Day 70
Day 85	Day 71 to Day 98
Day 113	Day 99 to Day 126
Day 141	Day 127 to Day 154
Day 169	Day 155 to Day 172
Day 176	Day 173 to Day 186
Day 197	Day 187 to Day 210
Day 225	Day 211 to End of Study

Table 6 Analysis Window Rules for Blood Histamine and Urinalysis

Visit Description	Window
Baseline	On or Prior to Day 1 Predose
Day 1	Predose on Day 1
Day 29	Day 1 Postdose to Day 42
Day 57	Day 43 to Day 70
Day 85	Day 71 to Day 98
Day 113	Day 99 to Day 126
Day 141	Day 127 to Day 158
Day 176	Day 159 to Day 186
Day 197	Day 187 to End of Study

In the event of multiple values from unscheduled or early termination assessments within a single analysis window, the value closest to the scheduled visit target study day will be used for analyses. If 2 values tie as closest to the time point (for example, 1 value is before and the other value is after the time point), then the later value will be selected. Data collected at all visits will be included in the data listings with visit presented as reported by the site.

5.3 Adjustment for Covariates

Efficacy analyses will be adjusted for baseline values and randomization stratum using ANCOVA, MMRM, or Cochran-Mantel-Haenszel tests, where applicable.

5.4 Handling of Dropouts, Missing Data, and Data Discrepancies

5.4.1 Missing Tissue Eosinophil Count at Week 24

Eosinophil counts will be set to missing if they are collected after subjects have experienced ICE. Subject with missing Week 24 tissue eosinophil count for any reason will be imputed with the eosinophil count collected from the postbaseline unscheduled visit prior to missing. If no eosinophil count is collected postbaseline, the subject will be considered treatment failure for tissue eosinophil response.

5.4.2 Missing Daily PRO Scores at Weeks 23–24

PRO daily scores will be set to missing if they are collected after subjects have experienced ICE. For the endpoints of change in the weekly average of PRO TSS, when ≥4 of 7 daily scores are available, the weekly average score will be calculated using the available daily scores. (See Appendix 1 for justification of allowing weekly average being calculated from ≥4 daily scores.) This calculation implies the missing daily scores are the same as the mean of the non-missing daily scores. When ≥4 daily scores are missing, the weekly score will be set to missing. For the calculation of the average PRO TSS over Weeks 23-24, if ≥1 weekly score is missing, the biweekly average of Weeks 23–24 will be considered missing.

Missing average of Weeks 23–24 TSS will be imputed using the Markov Chain Monte Carlo (MCMC) method. Baseline tissue eosinophil count and PRO TSS, and disease stratum (EG ± EoD vs. EoD without EG) will be included along with the average biweekly PRO TSS (W01_02, W03_04, W05_06, …, W21_22, W23_24) in the multivariate distribution construction. The imputation will be carried out for each treatment group separately and will be executed multiple (50) times (MI). The purpose of MI is to account for the imputation variability in the parameter estimates. An example of the SAS code for this imputation follows.

```
proc mi data=TSS seed=1357986420 nimpute=50
out=TSS_IMMPUTED ;
      by TRTMT ;
      class DISEASE_STRAT ;
      mcmc chain=multiple impute=full initial=em nbiter=200 niter=100 ;
var B_EOS DISEASE_STRAT (3 levels) B_TSS TSS_W01_02 TSS_W03_04
... TSS_W21_22 TSS_W23_24 ;
run ;
```

After the imputation, the least squares mean (LSM) and standard error (SE) are derived from the conventional ANCOVA from each imputed data set.

These LSM and SE are synthesized to derive the imputation variability-adjusted LSM and SE for each treatment group and for the between treatment difference. These synthesized LSM and SE are then used in the hypothesis test for the between treatment difference. An example of the SAS code for the synthesized LSM comparisons between treatment is as follows.

```
ods output ParameterEstimates=LSMDIF ;
proc mianalyze data=DIFF alpha=0.05 ;
      modeleffects LSM ;
      stderr SE ;
run ;
```

5.4.3 Missing or Partial Dates of Adverse Events or Concomitant Medications

Adverse events with incomplete start or stop dates (i.e., either day or month is missing) will be considered treatment-emergent adverse events (TEAE) unless the partial start date or the stop date confirms the AE started or ended prior to Study Day 1 (e.g., the day of the AE start date is unknown but the month and year indicate that the AE starts prior to Study Day 1). Adverse events with missing relationship to study drug will be included in the "Related" category for the summary tables and footnoted. Adverse events with missing severity will not be included in the by-severity summary table but will be footnoted in the table.

When there is incomplete information regarding dosing dates for prior and concomitant medication, the medication will be considered as a concomitant medication unless it contradicts with the stop date. For example, a medication will be considered a prior medication if the month and year of the end date indicates a date before Study Day 1 even though the start date could be missing.

5.5 Interim Analysis

No interim analysis is planned.

5.6 Timing of Data Analyses

There will be 2 database locks for this study:

- A provisional database lock after all subjects complete the Day 169 visit to allow for the analysis of safety and efficacy data through Day 169.
- A final database lock after all subjects complete the study to allow for the analysis of any safety data collected after Day 169.

For each database lock, applicable EDC data will be locked in order to protect write access after the following preconditions are fulfilled:

- All records are entered in the database.
- All AE are coded to the satisfaction of the Chief Medical Officer.
- All medications are coded to the satisfaction of the Chief Medical Officer.
- All data queries have been resolved.
- All decisions have been made regarding all protocol violations and ITT population exclusions.
- Written authorizations to lock the database are obtained from Allakos Clinical Data Management and the Chief Medical Officer.

The randomization code for this study will not be revealed until the previous preconditions are fulfilled and documentation of the provisional database lock is complete. After the provisional database lock, the randomization code will be made available to a restricted number of individuals at Allakos who are involved in the data analysis. Data analysis will commence after the provisional data lock and the final database lock, respectively. In addition, the PK and ADA data may be locked and assessed separately.

5.7 Multicenter Study

Analysis will be performed pooling data across study sites.

The study will have approximately 60 sites. Analysis of site effect will be based on analysis centers. Analysis centers will be formed in such a way that there are sufficient subjects per pooled analysis center for the assessment of site effects. Sites from the following geographic clusters have been identified for pooling data across study sites:

- **Analysis Center 1**: Massachusetts, Connecticut, Rhode Island, New York, New Jersey, Pennsylvania, District of Columbia, Maryland, Maine, Vermont, Delaware, and New Hampshire
- **Analysis Center 2**: Virginia, West Virginia, South Carolina, Georgia, Mississippi, Florida, Alabama, Louisiana
- **Analysis Center 3**: North Carolina, Tennessee, Ohio, Michigan, Indiana, Kentucky
- **Analysis Center 4**: South Dakota, Arkansas, Minnesota, Wisconsin, North Dakota, Illinois, Missouri, Iowa, Texas, Oklahoma, Kansas, Nebraska, Wyoming, Colorado
- **Analysis Center 5**: Utah, New Mexico, Arizona, California, Oregon, Washington, Idaho, Hawaii, Alaska, Montana, Nevada

5.8 Multiple Comparisons/Multiplicity Adjustment

To control for the family-wise type-I error rate, the following testing procedures will be implemented:

Test each co-primary efficacy endpoint at 2-sided α=0.05 level.

- If $p \leq 0.05$ for both co-primary endpoints, then reject the null hypothesis that AK002 is no different from placebo and accept the alternative hypothesis that AK002 is superior to placebo in reducing tissue eosinophil count and TSS score.
- If $p \leq 0.05$ for both co-primary endpoints, the hypothesis tests for the secondary endpoints will proceed in the prespecified order.
- If $p > 0.05$ for either co-primary endpoint, the null hypothesis is accepted. No hypothesis test will be performed for the secondary endpoints.
- The issue of multiplicity of secondary endpoints will be handled by statistical testing of these outcomes in a hierarchical fashion. The order of secondary endpoints is described in Section 2.2. If at any point during the analysis of the secondary endpoints, the statistical test is not significant at 2-sided α=0.05 level, the hypothesis testing procedure will stop. AK002 will be deemed superior to placebo for all endpoints prior to the stop.

5.9 Examination of Subgroups

Key endpoints will be summarized by subgroup to assess the consistency of the treatment effect across subgroups. Subgroups to be considered are:

- Baseline TSS (<28, ≥28)
- Disease strata (EG ± EoD, EoD)
- Gender (Male, Female)
- Age (<65, ≥65)
- Race (White, Non-White)

6. Statistical Analysis

6.1 Analysis Populations

The population of "all screened subjects" comprises subjects who signed the informed consent (ICF).

6.1.1 Safety Population

The safety population comprises randomized subjects who have received at least 1 dose of the study drug.

6.1.2 Intent-to-Treat Population

The intent-to-treat (ITT) population is defined as subjects who were randomized to treatment.

6.1.3 Modified Intent-to-Treat Population

The modified intent-to-treat (MITT) population is defined as subjects who were randomized and received at least 1 dose of the study drug.

6.1.4 Per Protocol Analysis Population

The per protocol (PP) population will include the MITT population who have received at least 1 dose of study drug and did not have major protocol violations potentially interfering with the efficacy assessment. The PP exclusion criteria will be specified prior to the database lock and unblinding.

6.1.5 Primary Analysis Population

The MITT population is the primary analysis population for all efficacy and safety analyses. Analysis using the ITT population will be carried out if ≥5 randomized subjects do not receive study drug. The PP population will be used to evaluate robustness for the primary endpoints and select secondary endpoints when appropriate.

6.2 Disposition of Subjects

Subject demographics and reasons for screening failure will be summarized for screen-failed subjects. Subjects (n and %) who completed or discontinued from the study will be tabulated by treatment group and for both treatment groups combined. The primary reasons for study discontinuation will be included in the tabulation. The primary reasons may include, but are not limited to, any of the following:

- Subject withdrew consent
- Lost to follow-up
- Administrative reason
- Adverse event
- Investigator decision
- Failure to follow required study procedures
- Other

Subject disposition will be summarized for all enrolled subjects. Subject counts for the Safety, MITT, and PP populations will be included in the table. A data listing for subject disposition will be presented for all enrolled subjects.

6.3 Protocol Deviations

Protocol deviations will include, but are not limited to

- Non-compliance with scheduled study visit
- Non-compliance with study treatment
- Received prohibited medications (see Protocol Section 8.1)
- Non-compliance with study inclusion or exclusion criteria
- Non-compliance with study assessment procedures

Subjects with major protocol deviations will be listed. The listing will include a brief description of the deviation, deviation category, and if applicable, study day when deviation occurred along with other pertinent information. If warranted by the sample size, subjects (n and %) with major protocol deviations will be tabulated by treatment group and by deviation category.

Subjects who are excluded from the PP population will be listed with reasons for exclusion.

6.4 Demographics and Baseline Subject Characteristics

Descriptive statistics for subject characteristics and baseline values will be presented for all populations by treatment group and both treatment groups combined. Continuous variables will be summarized with n, mean, SD, and median. Categorical variables will be summarized with n and % of subjects for each category for the MITT and PP populations.

6.5 Baseline Disease Characteristics

Baseline disease characteristics including demographics, weight, blood eosinophils (≥500 cells/μL), IgE, history of atopic dermatitis, functional GI symptoms, history of EG and/or EoD, gastric eosinophils (per 5 hpf in subject with EG), duodenal eosinophils (per 3hpf in EoD), TSS total score at baseline, TSS (<28 vs. ≥28), PRO symptom scores, and baseline elevated esophagitis histologic criteria will be included in the subject data listing.

6.6 Medical History

Subject incidence (n and %) of medical history (and current medical condition before signing the informed consent) will be tabulated by the Medical Dictionary for Regulatory Activities (MedDRA), version 21.0) System Organ Class (SOC) and Preferred Term (PT).

6.7 Electrocardiogram

A listing of electrocardiograms (ECG) overall interpretation at screening visit will be provided.

6.8 Pregnancy Test

A listing of pregnancy test results will be provided.

6.9 Baseline Diet

A listing of baseline and on-study diet assessment and compliance will be provided.

6.10 Treatments

6.10.1 Treatment Compliance and Extent of Exposure

Summaries of treatment compliance and exposure to AK002 will be based on the safety population.

Duration of treatment exposure is defined as the total number of days a subject is exposed to the study treatment. This will be calculated for each subject by taking the difference between the date of last dose minus the date of first dose, plus 1 (date of last dose – date of first dose +1). Duration of exposure will be summarized using descriptive statistics (n, mean, standard deviation, median, minimum, and maximum).

Treatment compliance is defined as the number of infusions that subjects received divided by 6 (i.e., total number of infusions expected) × 100%. Treatment compliance will be summarized descriptively. In addition, an overall compliance that includes adherence to treatment administration schedule and any interruptions will be included in the data listing.

The treatment infusion information (length, volume, rate, and interruption) will be included in the subject data listing.

6.10.2 Prior, Concomitant, and Newly Initiated Medications

Prior medications and concomitant medications will be extracted from the Prior/Concomitant Medication CRF. Medications taken prior to Study Day 1 will be considered as prior medications; medications taken on or after Study Day 1 will be considered as concomitant medications; and newly initiated medication refers to any medication with a start date ≥ Study Day 1. We note that a Prior Medication may also be a Concomitant Medication if the start date is prior to Study Day 1 and the end date is on/after Study Day 1. Medications will be coded using the WHO Drug Dictionary (WHODD March 2018 release) for PT and Anatomical Therapeutic Chemical (ATC) classification.

Medications will be tabulated separately based on the Safety population. The number (n and %) of subjects taking at least 1 medication and the number (%) of subjects taking each medication at the preferred term level will be tabulated by ATC4, and PT. Subjects taking the same PT medication twice will only be counted once.

A subject data listing will be provided to include the reported medication name, the WHODD PT, ATC4, study day and pertinent subject information. A separate data listing will include subjects who have received prohibited medications.

6.11 Analysis of Primary Efficacy Endpoints

To assist the data interpretation, all ICE will be listed with pertinent subject information (subject ID, date of last infusion, ICE start/stop day, nature of ICE (reason for treatment/study discontinuation if applicable, description of AE, description of prohibited/restricted medication), baseline tissue eosinophil count and TSS, and most recent tissue eosinophil count and 2-week average of TSS). A summary table will be created to present number of subjects (n, %) by treatment group:

- Subjects with any ICE
  - Subjects by ICE category

6.11.1 Analysis of the First Co-Primary Endpoint

The first co-primary endpoint will be analyzed using the imputed data set (Section 5.4.1). Fisher's exact test (primary analysis) will be conducted comparing AK002 with placebo for the proportion of tissue eosinophil responders. Proportion of responders and the associated 95% confidence interval will be presented for each treatment group. The between group difference and the associated 95% confidence interval will also be computed and presented. A sample SAS code is as follows.

```
* COMPUTE 95% EXACT CONFIDENCE INTERVAL FOR %RESPONSE FOR INDIVIDUAL
TREATMENT ;
ods output BinomialCLs=CL ;
proc freq data=ADEF ;
      table RESP / out=CNTS bin(cl=midp) ;
      by TRTAN ;
run;

* COMPUTE 95% EXACT CONFIDENCE INTERVAL AND FISHERS EXACT P-VALUE FOR
BETWEEN TREATMENT DIFFERENCE ;
ods output FishersExact=PVAL(where=(name1='XP2_FISH'))
RiskDiffCol2=DIFF(where=(row='Difference') ;
proc freq data=ADEF ;
      table TRTPN*RESP / riskdiff(cl=exact) exact  ;
run ;
```

6.11.2 Sensitivity Analysis of the First Co-Primary Endpoint

Sensitivity analysis may be carried out using the Cochran-Mantel-Haenszel (CMH) test stratified by the randomization stratification factors (baseline TSS [<28 vs. ≥28] and Disease Strata [EG ± EoD vs. EoD]) to assess robustness of the Fisher's exact test results.

A sample SAS code for the pairwise comparison is as follows.

```
* COMPUTES CMH P-VALUE AND NEWCOMBE COMMON RISK DIFFERENCE ;
ods output CMH=PVAL(where=(althypothesis='Row Mean Scores Differ'))
CommonPdiff=DIFF;
proc freq data=ADEF ;
      tables TSS_STRAT*DISEASE_STRAT*TRTP*RESP / cmh commonriskdiff
(cl=NEWCOMBE) ;
run ;
```

The rationale for specifying the Fisher's exact test as the primary analysis as opposed to specifying the CMH is because when a stratum has 100% response for 1 treatment group and 0% response for another treatment group, the CMH test may lose efficiency or not be computable. This is evident from the Phase 2 study outcome.

6.11.3 Subgroup Analysis of the First Co-primary Endpoint

Analysis comparing AK002 and placebo will use the Fisher's exact test for each of the subgroups defined in Section 5.9.

6.11.4 Analysis of the Second Co-Primary Endpoint

The second co-primary endpoint will be analyzed by ANCOVA (primary analysis) using the imputed data set (Section 5.4.2). The LS mean, SE, and 95% CI for each treatment group and for the between group difference will be derived from ANCOVA with treatment as factor, baseline PRO TSS (continuous) and disease strata (EG ± EoD, EoD) as covariates. The hypothesis test for the treatment effect will be carried out by the F-test. The synthesizing method will be used to combine the results from multiple imputations.

6.11.5 Sensitivity Analyses of the Second Co-Primary Endpoint

The two sensitivity analyses planned below are constructed based on the already imputed biweekly data set from Section 5.4.2. From this imputed data set, the biweekly TSS is set to missing if it is derived from the daily scores containing ≥4 daily scores collected on/after the ICE. This setting creates the monotone missing data set.

The first sensitivity analysis is based on the placebo-based pattern-mixture model for the missing data imputation under the missing not at random (MNAR) assumption. In this model, subjects

from the active treatment group after the ICE are assumed to behave like the subjects from the placebo group. Their missing data are imputed using the response profile from the placebo subjects who have similar baseline covariates and prior response trajectory. The sample SAS code follows.

```
proc mi seed=135791 nimpute=10;
      class TRTMT;
      var B_EOS DISEASE_STRAT B_TSS TSS_W01_02 TSS_W03_04 ...
      TSS_W21_22
            TSS_W23_24;
      monotone reg;
      mnar model(TSS_W01_02 TSS_W03_04 ... TSS_W21_22 TSS_W23_24 /
            modelobs=(TRTMT='0'));  * CODE 0 IS FOR PLACEBO GROUP ;
run ;
```

The second sensitivity analysis is based on the tipping point method. In this method, the missing biweekly TSS will be imputed with different adjustments for the active treatment subjects and placebo subjects under the MNAR assumption in which the search for a tipping point reverses the study conclusion (i.e., p-value no longer <0.05 for the treatment effect). A shift ranging from *-x* to *+y* representing different levels of improvement or worsening due to ICE will be added to the imputed TSS. An example of the SAS code is as follows.

```
proc mi seed=579864 nimpute=10;
      class DISEASE_STRAT TRTMT;
      var B_EOS DISEASE_STRAT B_TSS TSS_W01_02 TSS_W03_04 ...
            TSS_W21_22 TSS_W23_24;
      monotone reg;
      mnar  adjust(TSS_W01_02 / adjustobs=(TRTMT='1') shift=&d1.)
            adjust(TSS_W03_04 / adjustobs=(TRTMT='1') shift=&d1.)
            . . .
            adjust(TSS_W21_22 / adjustobs=(TRTMT='1') shift=&d1.)
            adjust(TSS_W23_24 / adjustobs=(TRTMT='1') shift=&d1.)
            adjust(TSS_W01_02 / adjustobs=(TRTMT='0') shift=&d0.)
            adjust(TSS_W03_04 / adjustobs=(TRTMT='0') shift=&d0.)
            . . .
            adjust(TSS_W21_22 / adjustobs=(TRTMT='0') shift=&d0.)
            adjust(TSS_W23_24 / adjustobs=(TRTMT='0') shift=&d0.);
run;
```

Note, a positive adjustment (shift>0) will increase (worsen) TSS and a negative adjustment (shift<0) will reduce (improve) TSS. Varying shift parameter from 0 (minimum) to maximum, where maximum shift parameter will be guided by the mean difference between subjects that drop out after the t^{th} visit and subjects that continue. This sensitivity analysis will vary the shift parameter independently between the active treatment subjects and placebo subjects.

After the data are imputed, the ANCOVA with the synthesizing method will be applied using methods described in Section 5.4.2.

P-values for the between treatment comparison from the synthesized ANCOVA results will be plotted against the shift parameters to demonstrate the robustness of the study conclusion against the various imputation strategies to account for the ICE effect.

6.11.6 Supplementary Analyses

Analysis of Section 6.11.2 will be repeated for the Per Protocol population.

Additional analysis will investigate the treatment effect on PRO TSS change from baseline in the subgroups defined in Section 5.9. This analysis will use ANCOVA similarly to the primary analysis (Section 6.11.2).

Empirical cumulative distribution function (eCDF) of the change from baseline in TSS will be plotted by treatment group to demonstrate consistency of the treatment effect.

6.12 Analysis of Secondary Efficacy Endpoints

6.12.1 Change in Tissue Eosinophil Count

For subjects who provide gastric or duodenal only biopsy, the calculation will be based on the average count of the highest readings from the respective mucosa at baseline and Day 169 (Week 24). For subjects who provide both gastric and duodenal biopsies, the calculation will be based on the average count of the highest readings from the Day 169 biopsy that corresponds in location (gastric or duodenal) to the location with the highest average count at baseline. The change in tissue eosinophil count from baseline to Day 169 will be analyzed using ANCOVA with treatment as factor, baseline eosinophil counts, baseline PRO TSS, and disease status as covariates. The LSM, SE, and 95% CI for individual treatment groups, and LSM, SE, 95% CI, and p-value for the between treatment difference will be presented.

6.12.2 Proportion of Subjects Achieving Peak Gastric and/or Duodenal Intraepithelial Eosinophil Count of ≤1 Cell/hpf

This endpoint will be analyzed using the Fisher's exact test similar to the analysis for the first co-primary endpoint (Section 6.11.1).

6.12.3 Proportion of Treatment Responders

Treatment responder is defined as >30% improvement in TSS and mean eosinophil count ≤4 cells/hpf in 5 highest gastric hpf and/or mean eosinophil count ≤15 cells/hpf in 3 highest duodenal hpf. This endpoint will be analyzed using Fisher's exact test similar to Section 6.11.1.

6.12.4 Proportion of Subjects with ≥50% Reduction and ≥70% Reduction in TSS from Baseline to Weeks 23–24

These 2 endpoints will be analyzed using the CMH test stratified by the randomization stratification factors.

6.12.5 Change in Weekly TSS Over Time

Prior to the analysis, the missing weekly TSS scores will be imputed similarly to the imputation for the missing biweekly PRO TSS scores (Section 5.4.2). The imputed data sets will be analyzed using the mixed model for repeat measures ([redacted]) with treatment, week, and treatment-by-week interaction as fixed factors, baseline PRO TSS (continuous), and EoD without EG (Y/N) as covariates, and subject as repeated measure unit over time. The model will use change from baseline in the weekly TSS from Weeks 1 through Week 24.

The model variance-covariance matrix will be unstructured. If the computation does not converge, the covariance matrix will take the form of Toeplitz, AR(1), and compound symmetry, whichever converges first. The improved Kenward-Rodger's method (SAS KR2 option) will be used to derive the denominator degrees-of-freedom. The LSM and 95% CI for the between group difference will be estimated using the simple contrast at each week.

A sample SAS code is provided as follows.

```
proc mixed data=ADEF method=REML ;
   class TRTPN USUBJID WEEK ;
   model CHG = B_TSS DISEASE_STRAT TRTPN WEEK TRTPN*WEEK / ddfm=KR2 ;
   repeated WEEK / subject=USUBJID(TRTPN) type=UN ;
   lsmeans TRTPN*WEEK / pdiff cl ;
run ;
```

The rationale for using the [redacted] model is that this analysis is appropriate assuming subjects with missing data behave similarly to other subjects in the same treatment group. Having included in the model baseline symptom score (TSS), disease status (EG ± EoD vs. EoD), and response trajectory prior to missing, it is reasonably to believe that the missing data are either completely at random (MCAR) or at random (MAR) in this setting. In another words, if missing mechanism is not MCAR, it is likely explainable by the above model covariates and response trajectory. A sample SAS code is included in Appendix 4.

6.13 Analysis of Exploratory Endpoints

6.13.1 Change in [redacted]

Change from baseline in the [redacted]

CCI This analysis method is similar to the analysis for the second co-primary endpoint (Section 6.11.2).

6.13.2 Change in CCI

Change from baseline in CCI summary (including domain scores) will be analyzed using CCI similar to Section 6.12.5.

6.13.3 Subjects Achieving CCI at Week 24

For subjects with CCI at baseline, the proportion of subjects achieving CCI at Week 24 will be analyzed using the CCI similar to Section 6.11.1.

6.13.4 Change in CCI

Change from baseline in CCI similar to Section 6.12.5 with CCI

6.13.5 CCI

Change from baseline in CCI as described in Section 6.12.5. The model will include CCI

- CCI
- CCI
- CCI
- CCI
- CCI
- CCI

- CCI [redacted]
- CCI [redacted]

6.13.6 CCI [redacted]

Methods for CCI [redacted] and CCI [redacted] are described in a separate analysis plan.

6.14 Analysis of Pharmacokinetic Endpoint

The analysis of the PK concentration will be based on the Safety population. AK002 concentrations will be listed by treatment, subject, nominal time, and actual time. Concentrations that are below limit of quantification (BLOQ) will be indicated in the listing.

Plasma concentrations will be summarized at each nominal time point separately for each treatment. The following descriptive statistics will be presented: n, arithmetic mean, SD, geometric mean, % CV, median, minimum, and maximum.

Individual plasma concentration vs. actual times will be plotted for each subject in linear and semi-logarithmic scales, placed on the same page. Mean plasma concentration at the scheduled time points will be plotted for each treatment in linear and semi-logarithmic scale, with the associated standard errors (for linear scale only) at each scheduled time point.

In the plot, concentrations that are BLOQ will be assigned a value of 0 if they are collected predose or assigned a value of 1/2 the lower limit of quantification (LLOQ) if they are collected postdose.

Analysis of PK parameters are specified separately in a PK analysis plan.

6.15 Safety Analyses

6.15.1 Adverse Events

Safety assessments will be based mainly on the nature, frequency, relationship, and severity of adverse events (AE). The AE will be coded by primary SOC and PT according to MedDRA (version 21.0). The treatment-emergent adverse events (TEAE) will be summarized by the number and percentage (n and %) of subjects in each SOC and PT.

For summaries by relationship to study drug, "possibly related" will be combined with "related", and "unlikely/remotely related" will be combined with "not related." When multiple AE are reported with the same PT, the AE of the strongest relation to study drug will be included in the summary by relationship, and the AE of the most severe grade will be included in the summary by severity table.

The following AE incidence tables will be presented.

- Overview of TEAE to include
  - Number (%) of subjects who reported at least 1 TEAE
  - Number (%) of subjects who reported at least 1 treatment-related TEAE
  - Number (%) of subjects who reported at least 1 severe TEAE
  - Number (%) of subjects who reported at least 1 serious TEAE
  - Number (%) of subjects who reported at least 1 TEAE leading to treatment discontinuation
  - Number (%) of subjects who reported at least 1 TEAE leading to study discontinuation
  - Number (%) of subjects who reported at least 1 TEAE of special interest (TEAESI)
- TEAE by PT sorted by decreasing order of subject incidence in the combined treatment group
- TEAE by SOC and PT in alphabetical order
- TEAE by SOC, PT, and maximum severity
- TEAE by SOC, PT, and strongest relationship to study drug
- TEAE leading to treatment discontinuation by SOC and PT, if warranted by sample size
- TEAE leading to study discontinuation by SOC and PT, if warranted by sample size
- Serious TEAE by SOC and PT, if warranted by sample size
- TEAESI by SOC and PT

All AE will be listed with onset/stop day, relationship to study drug, severity, action taken, and outcome. Pertinent subject information including treatment group and demographics will also be included.

Separate listings will be provided for TEAE leading to study discontinuation, TEAE of special interest, and treatment-emergent serious AE (TESAE).

6.15.2 Laboratory Test

Laboratory data will be summarized by presenting summary statistics of raw data and change from baseline values (means, medians, standard deviations, ranges). For quantitative laboratory parameters, both actual values and change from baseline values will be summarized at each visit.

Shift tables will be presented, in which, lab test results at baseline and post-baseline visit will be classified into below (<LLN), within (≥LLN and ≤ULN), and above (>ULN) normal ranges. Subject incidences (n and %) will be presented for the shift from baseline to the post-baseline visits.

Note that the analysis window will be applied for the visits.

A complete laboratory data listing, including hematology, biochemistry, and urinalysis will be provided for all subjects.

6.15.3 Vital Signs, Height and Weight, and Other Safety Measures

Vital signs will be summarized by presenting summary statistics of raw data and change from baseline values (means, medians, standard deviations, ranges) for each visit (per analysis window) and time point. A data listing will include vital signs from all visits.

6.15.4 Electrocardiogram

Incidence of the Investigator's overall assessment (normal, abnormal – not clinically significant, and abnormal – clinically significant) will be summarized.

6.15.5 Physical Examination

New or worsening symptoms in the symptom-directed physical exams will be included in the by-subject data listing.

6.15.6 Analysis of Anti-Drug Antibody

A data listing of anti-drug-antibodies (ADA) results will be provided for all subjects. Number (%) of subjects who are confirmed ADA-positive at any time after receiving study drug and number (%) of subjects who are confirmed ADA-positive at the end of study will be cross-tabulated by their ADA status and titers at predose.

7. Validation

The Clinical Operations and Data Management at Allakos will work with the EDC/Data Management (DM) vendor to ensure that the data collected for the study are of the highest quality possible. The study monitor will be responsible for reviewing and verifying the accuracy

of the data recorded on the eCRF directly from source documents at the investigative site. The DM vendor will be responsible for performing edit checks and reviewing all data entered into the electronic database to identify discrepant and/or inconsistent values and to send queries to the clinical sites. The Investigator will be responsible for answering queries about discrepant data and providing electronic signatures to confirm data integrity.

The programming of tables, listings and figures (TLF) based on the clinical data is outsourced. Allakos seeks to ensure the quality of the reports provided by the CRO in the form of TLF passing a rigorous validation process as follows:

- Derived datasets will be independently reprogrammed by a second programmer. The separate datasets produced by the 2 programmers must match 100%.
- Tables will be independently reprogrammed by a second programmer, and the results from both programs must match.
- Figures will be checked for consistency against corresponding tables and listings or independently reprogrammed if there are no corresponding tables or listings.
- Listings will be checked for consistency against corresponding tables, figures, and derived datasets.

All post hoc analysis, if accepted for inclusion in the study report, will follow the same quality process as stated above. The entire set of TLF will be checked for completeness and consistency prior to delivery to Allakos.

8. References

International Council for Harmonisation (ICH). Guideline for industry E3, structure and content of clinical study reports, July 1996.

Food and Drug Administration. E9 (R1) Statistical Principles for clinical trials: addendum: estimands and sensitivity analysis in clinical trials. Guidance for industry, ICH, May 2021.

9. Appendices

9.1 Appendix 1: Effect of Missing Data on the Weekly Mean of TSS

Following the FDA's suggestion for the evaluation of the missing daily diary effect on the weekly mean of TSS score, data from the Phase 2 study are extracted. The Phase 2 study has diaries collected daily over 14 weeks of the treatment period. A simulation study was conducted examining the changes in standard error by the number of missing daily diaries that were randomly created. The step are as follows.

1) A complete date set is created in which a subject week is kept only if the week has all 7 daily diaries. This data set is at the subject-day level.

2) For each week in the complete data set, 1 of the 7 daily diaries is randomly set to missing. This creates a data set with 1 day per week missing.

3) Step 2 is repeated for 2 of the 7 daily diaries are randomly set to missing to create a data set with 2 day per week missing.

   Finally, a data set is created with 6 days per week are randomly set to missing.

4) For each of the 7 data sets (0 day/week missing, 1 day/week missing, …, 6 days/week missing), daily diaries are averaged to derive the weekly mean for each subject at each week.

5) Weekly means of the diaries are further averaged across all subjects to derive the population means and population standard error of the diaries for each week.

6) Steps 2 through 5 are repeated 500 times. Each time, the daily diaries are randomly set to missing.

7) The 500 weekly population standard errors from Step 6 are averaged and presented in the table below.

As expected, the standard error increases as the number of missing daily diaries increases, indicating the estimate of weekly TSS mean becomes less accurate. However, the increase in the standard error does not exceed 2% when allowing 3 out of 7 daily diaries missing, and does not exceed 4% when allowing 4 out of 7 daily diaries missing.

Table 7 Standard Errors for the Mean Weekly TSS Diaries by Number of Daily Diaries Missing Averaged over 500 Simulations

Week	Number of Daily Diaries Missing						
	0 Day	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days
1	3.092	3.100	3.115	3.132	3.166	3.232	3.405
2	2.917	2.923	2.929	2.941	2.964	3.003	3.102
3	3.070	3.081	3.088	3.095	3.125	3.198	3.321
4	2.862	2.868	2.874	2.881	2.894	2.926	2.994
5	3.111	3.118	3.122	3.134	3.149	3.187	3.293
6	3.135	3.137	3.147	3.154	3.170	3.191	3.279
7	3.039	3.045	3.054	3.067	3.093	3.140	3.283
8	3.842	3.846	3.849	3.858	3.871	3.884	3.959
9	2.854	2.869	2.874	2.899	2.932	3.021	3.176
10	2.823	2.829	2.840	2.845	2.857	2.886	3.010
11	2.997	3.002	3.010	3.008	3.037	3.069	3.148
12	2.656	2.661	2.671	2.689	2.691	2.731	2.842
13	2.864	2.879	2.886	2.905	2.950	3.011	3.208
14	2.388	2.402	2.414	2.437	2.474	2.580	2.795

Source: K:\AK002-016\planning\MissingOnSE.sas 08MAR2020

9.2 Appendix 2: Intercurrent Events

A. Use of any of the following prohibited medications during the course of the study as described

1) Immunosuppressive or immunomodulatory drugs (e.g., IL-5 modulators, i.e., benralizumab, reslizumab, mepolizumab; IL-4 and IL-13 antagonists, i.e., dupilumab; calcineurin inhibitors, i.e., cyclosporin, tacrolimus;, mTOR inhibitors, i.e., sirolimus, everolimus; anti-metabolites, i.e., azathioprine, methotrexate, 6-mercaptopurine, leflunomide, mycophenolate mofetil; alkylating agents, i.e., cyclophosphamide; TNF inhibitors, i.e., infliximab, adalimumab; anti-IgE antibodies, i.e., omalizumab; and eosinophil-depleting drugs, i.e., pramipexole.

a) Use at any point during the study.

2) Glucocorticoids

a) Initiation of any course of treatment or single use of systemic or swallowed corticosteroids at a dose of >10 mg/day prednisone or equivalent starting at Week 21 and through the end of Week 24. This does not include use of corticosteroids given as pre-infusion prophylaxis or treatment of IRR.

B. Use of restricted medications outside the protocol-defined specifications

1) Glucocorticoids

a) Discontinuation of a previously stable dose of systemic or swallowed corticosteroids (≤10 mg/day prednisone or equivalent) after the screening period.

b) Initiation of a new course of systemic or swallowed corticosteroids (≤10 mg/day prednisone or equivalent) following the screening period and continuing to at least study Week 21 (within 2 weeks of efficacy assessment at Weeks 23 and 24).

2) Proton Pump Inhibitors

a) Discontinuation of a previously stable dose of an oral PPI after the screening period and before the end of Week 24.

b) Initiation or increase in the dose of an oral PPI following the screening period and continuing to at least study Week 21 (within 2 weeks of efficacy assessment at Weeks 23 and 24).

3) Sodium cromolyn

a) Discontinuation of a previously stable dose of sodium cromolyn after the screening period and before the end of Week 24

b) Initiation or increase in the dose of sodium cromolyn following the screening period and continuing to at least study Week 22 (within 1 week of efficacy assessment at Weeks 23 and 24).

C. Discontinuation of Investigational Medical Product

1) Discontinuation of study agent due to:

a) Infusion-related reaction

b) Subject withdrawal

c) Meeting study withdrawal criteria (Protocol Section 13.10)

Per protocol, subjects are required to remain on stable doses of either PPI or corticosteroids (≤10 mg prednisone equivalent per day) throughout the course of the study if they are on a stable dose during the screening period. Because of the potential impact of corticosteroids and PPI on both histologic assessment of tissue eosinophil levels as well as subject reported GI symptoms, subjects must meet the entry requirements for both histology and symptom burden (TSS6) if they are on any of these medications at screening and therefore must remain on a stable dose throughout the study in order to meaningfully interpret any change from baseline during the evaluation phase. Conversely, subjects starting these medications due to unforeseen medical necessity following the screening period and remaining on them through a time point sufficiently close to the evaluation period to potentially impact assessment of symptoms and/or histologic assessments will be considered as intercurrent events. For PPI and corticosteroids (≤10 mg/day prednisone or equivalent) a conservative washout period is estimated at 2 weeks, and therefore, continuation of a new course of therapy beyond Week 21 could be considered as a potential ICE.

Sodium cromolyn is a mast cell stabilizer and could potentially impact reported symptoms in a similar way to 1 of the mechanisms of action of AK002. As such, discontinuation of a previously stable dose of sodium cromolyn after the screening period and before the end of Week 24 could be considered as a potential ICE. Likewise, the initiation or increase in the dose of sodium cromolyn following the screening period and continuing to at least study Week 22 (within 1 week of efficacy assessment at Weeks 23 and 24) could be considered as a potential ICE. It is estimated that 1 week would be a sufficient washout period for the effects of sodium cromolyn.

Antihistamines are unlikely to impact histologic assessments in the study but could have slight impact in some symptoms depending on the type, dose, and duration of use. As such, subjects are required to stay on stable doses of antihistamines if they are taking them at the time of screening and to remain on them through completion of the study. Due to unforeseen medical circumstances and acknowledging that some subjects may have concurrent allergic conditions, the use of these medications by some subjects during the study is anticipated.

9.3 Appendix 3: List of Tables, Figures, and Listings

Tables

Table 14.1.13	Baseline Subject Reported Outcome Total and Symptom Scores (MITT Population)
Table 14.1.14	Baseline Vital Signs (MITT Population)
Table 14.1.15	Subject Incidence of Intercurrent Events (MITT Population)
Table 14.2.1.1	Tissue Eosinophil Responder at Week 24 – Primary and Sensitivity Analysis (MITT Population)
Table 14.2.1.2.1	Tissue Eosinophil Responder at Week 24 by Disease Subgroup (MITT Population)
Table 14.2.1.2.2	Tissue Eosinophil Responder at Week 24 by Gender (MITT Population)
Table 14.2.1.2.3	Tissue Eosinophil Responder at Week 24 by Age Subgroup (MITT Population)
Table 14.2.1.2.4	Tissue Eosinophil Responder at Week 24 by Baseline Peripheral Eosinophil Count Subgroup (MITT Population)
Table 14.2.2.1	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores – Primary Analysis – Multiple Imputation for Missing Weekly TSS (MITT Population)
Table 14.2.2.2	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores – Sensitivity Analysis 1 – Weekly TSS Due to ICE are Replaced/Imputed Using Placebo-Based Pattern-Mixture Model Multiple Imputation (MITT Population)
Table 14.2.2.3	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores – Sensitivity Analysis 2 – Weekly TSS Due to ICE are Replaced/Imputed Using Multiple Imputation by the Tipping Point Method (MITT Population)
Table 14.2.2.4	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores – Multiple Imputation (Per Protocol Population)
Table 14.2.2.5.1	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores by Disease Subgroup – Multiple Imputation (MITT Population)
Table 14.2.2.5.2	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores by Gender – Multiple Imputation (MITT Population)

Table 14.2.2.5.3	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores by Age Subgroup – Multiple Imputation (MITT Population)
Table 14.2.2.5.4	Change from Baseline in Average of Weeks 23-24 PRO Total Symptom Scores by Baseline Peripheral Eosinophil Count Subgroup – Multiple Imputation (MITT Population)
Table 14.2.3	Change from Baseline to Week 24 Tissue Eosinophil Count (MITT Population)
Table 14.2.4	Subjects Achieving Peak Gastric and/or Duodenal Intraepithelial Eosinophils Count ≤1 cell/hpf (MITT Population)
Table 14.2.5	Subjects Achieving Tissue Eosinophil and TSS Treatment Response (MITT Population)
Table 14.2.6	Subjects with ≥50% Reduction and ≥70% Reduction in TSS from Baseline to Weeks 23-24
Table 14.2.7	Change from Baseline in Weekly TSS from Week 1 Through Week 24 Using MMRM (MITT Population)
Table 14.2.8	Change from Baseline to Week 24 Tissue CCI [redacted] (MITT Population)
Table 14.2.9.1	Change from Baseline in CCI [redacted] – Physical Component Summary (MITT Population)
Table 14.2.9.2	Change from Baseline in CCI [redacted] – Mental Component Summary (MITT Population)
Table 14.2.10	Subjects Achieving Week 24 CCI [redacted] (MITT Subjects with Concomitant EoE)
Table 14.2.11	Change from Baseline in CCI [redacted] (MITT Population)
Table 14.2.12.1	Change from Baseline in Weekly CCI [redacted] (MITT Population)
Table 14.2.12.2	Change from Baseline in Weekly CCI [redacted] (MITT Population)
Table 14.2.12.3a	Change from Baseline in Weekly CCI [redacted] (MITT Population)

Table 14.2.12.3b	Change from Baseline in Weekly [redacted] (MITT Population)
Table 14.2.12.4a	Change from Baseline in Weekly [redacted] (MITT Population)
Table 14.2.12.4b	Change from Baseline in Weekly [redacted] (MITT Population)
Table 14.2.12.5	Change from Baseline in Weekly [redacted] (MITT Population)
Table 14.2.12.6	Change from Baseline in Weekly [redacted] (MITT Population)
Table 14.2.12.7	Change from Baseline in Weekly [redacted] (MITT Population)
Table 14.2.12.8	Change from Baseline in Weekly [redacted] (MITT Population)
Table 14.2.13	Severity of Dysphagia at Baseline and Post-Baseline (MITT Population)
Table 14.2.14	Morphological Assessment of Gastric Biopsies by Sydney System (MITT Population)
Table 14.2.15	Morphological Assessment of Duodenal Biopsies by Marsh Scale Classification (MITT Population)
Table 14.2.17	Change from Baseline in Body Weight (MITT Population)
Table 14.2.18	Total Serum IgE Concentration (Per Protocol Population)
Table 14.2.19	Subject Incidence of Positive Anti-Drug (AK002) Antibody (ADA) Test (Safety Population)
Table 14.3.1.1	Overview of Treatment-Emergent Adverse Events (Safety Population)
Table 14.3.1.2	Treatment-Emergent Adverse Events by Preferred Term: Subject Incidence (Safety Population)
Table 14.3.1.3	Treatment-Emergent Adverse Events by System Organ Class and Preferred Term: Subject Incidence (Safety Population)

Figures

Listings

Listing 16.1.7	Subject Randomization Schedule (Randomized Population)
Listing 16.2.1.1	Subjects Disposition and Reason for Disposition (Randomized Population)
Listing 16.2.1.2	Subjects Who had Major Protocol Deviations (Randomized Population)
Listing 16.2.1.3	Subjects Inclusion and the Reason for Exclusion from Analysis Populations (Randomized Population)
Listing 16.2.1.4.1	Demographic and Baseline Characteristics (Safety Population)
Listing 16.2.1.4.2	Demographic and Baseline Characteristics – History of Disease (Safety Population)
Listing 16.2.1.5	Medical History (Safety Population)
Listing 16.2.1.6	Diet Assessment (Safety Population)
Listing 16.2.1.7.1	Prior Medications (Safety Population)
Listing 16.2.1.7.2	Concomitant Medications (Safety Population)
Listing 16.2.1.8	Intercurrent Events (Safety Population)
Listing 16.2.2.1.1	Esophago-Gastro-Duodenoscopy (EGD) Biopsy Finding (Safety Population)
Listing 16.2.2.1.2	Duodenal biopsy assessments (Safety Population)
Listing 16.2.2.1.3	Esophageal biopsy assessments (Safety Population)
Listing 16.2.2.2a	Eosinophilic Gastritis and Gastroenteritis Questionnaire – Question Text (Safety Population)
Listing 16.2.2.2b	Eosinophilic Gastritis and Gastroenteritis Questionnaire – Daily Score (Safety Population)
Listing 16.2.2.2c	Eosinophilic Gastritis and Gastroenteritis Questionnaire – Weekly Score (Safety Population)
Listing 16.2.2.3	Eosinophilic Gastritis Endoscopic Reference Score (EGERS) (Safety Population)
Listing 16.2.2.4a	[redacted] Questionnaire – Question Text

Listing 16.2.2.4b	[redacted] Questionnaire – Individual Questions (ITT Population)
Listing 16.2.2.4c	[redacted] Questionnaire – Domain Scores (ITT Population)
Listing 16.2.2.5a	Intensity of Symptoms Questionnaire – Question Text (Safety Population)
Listing 16.2.2.5b	Intensity of Symptoms Questionnaire – Individual Scores (Safety Population)
Listing 16.2.3.1.1	All Adverse Events (Safety Population)
Listing 16.2.3.1.2	Serious Adverse Events (Safety Population)
Listing 16.2.3.1.3	Adverse Events with Outcome of Death (Safety Population)
Listing 16.2.3.1.4	Adverse Events Leading to Study Drug Discontinuation (Safety Population)
Listing 16.2.3.1.5	Adverse Events of Special Interest (Safety Population)
Listing 16.2.3.2.1	Clinical Laboratory Tests-Hematology (Safety Population)
Listing 16.2.3.2.2	Clinical Laboratory Tests – Blood Chemistry (Safety Population)
Listing 16.2.3.2.3	Clinical Laboratory Tests – Urinalysis (Safety Population)
Listing 16.2.3.2.4	Clinical Laboratory Tests – Diagnostic Immunology/Molecular Biology (Safety Population)
Listing 16.2.3.2.5	Clinical Laboratory Tests – Endocrinology (Safety Population)
Listing 16.2.3.2.6	Clinical Laboratory Tests – Parasitology (Safety Population)
Listing 16.2.3.2.7	Pre-Infusion Blood Collection (Safety Population)
Listing 16.2.3.3.1	Substance Use – Alcohol Consumption at Screening (Safety Population)
Listing 16.2.3.3.2	Urine Pregnancy Test (Safety Population)
Listing 16.2.3.4.1	Study Drug Administration (Safety Population)
Listing 16.2.3.4.2	Study Drug Accountability (Safety Population)
Listing 16.2.3.5	Vital Signs (Safety Population)
Listing 16.2.3.6	Subject Body Size (Safety Population)
Listing 16.2.3.7.1	12-Lead ECG Intervals at Baseline (Safety Population)

9.4 Appendix 4: Example SAS Codes

```
/** Multiple Imputation **/
proc mi data=TSS seed=1357986420 nimpute=50 out=TSS_IMPUTED1 ;
       by TRTMT ;
       class DISEASE_STRAT ;
       mcmc chain=multiple impute=full initial=em nbiter=200 niter=100 ;
var B_EOS DISEASE_STRAT B_TSS TSS_W01_02 TSS_W03_04... TSS_W21_22 TSS_W23_24 ;
quit;
/** Recalculate change from baseline **/
data tss_imputed2;
   set tss_imputed1;
   chg=tss_w23_24 - b_tss;
run;

proc sort data=tss_imputed2 out=tss_imputed;
   by _imputation_;
run;

/** ANCOVA by imputation **/
ods output diffs=diffs lsmeans=lsm;
proc mixed data=tss_imputed method=reml;
   by _imputation_;
   class TRTMT;
   model CHG=DISEASE_STRAT B_TSS TRTMT;
   lsmeans TRTMT / diff;
quit;
ods output close;

/** Combine LS Means (by treatment and difference) **/
data lsmdiff;
   set lsm diffs(in=d);
   if d then trtmt='DIFF';
run;


proc sort;

   by TRTMT;
run;

/** Synthesized results **/
ods output ParameterEstimates=SYNDIF;
proc mianalyze data=LSMDIFF alpha=0.05;
    by TRTMT;
    modeleffects ESTIMATE;
    stderr;
run ;
ods output close;
```