STATISTICAL ANALYSIS PLAN

SP-1011-002

FLUTICASONE IN EOSINOPHILIC ESOPHAGITIS (FLUTE): A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, DOSE RANGING, AND MAINTENANCE STUDY OF APT-1011 IN SUBJECTS WITH EOSINOPHILIC ESOPHAGITIS

AUTHOR: [redacted]

VERSION NUMBER AND DATE: V3.0, 06NOV2019

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan V3.0 (Dated 06NOV2019) for Protocol SP-1011-002

	Name	Signature	Date
Author:	[redacted]		
Position:	Statistical Team Lead		
Company:	IQVIA		

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

	Name	Signature	Date
Approved By:	[redacted]		
Position:	Senior Reviewer		
Company:	IQVIA		
Approved By:	[redacted]		
Position:	Statistical Team Lead		
Company:	IQVIA		
Approved By:	[redacted]		
Position:	Director of Clinical Operations for Adare Pharmaceuticals		
Company:	Adare		

Approved By:	[redacted]		
Position:	Statistical Team Lead for Adare Pharmaceuticals		
Company:	Adare		

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
0.1	18APR2018 06JUN2018	[redacted] [redacted]	Not Applicable – First Version
2.0	05JAN2018	[redacted]	Added FAS population
3.0	06NOV2019	[redacted]	Updated for Final Analysis

TABLE OF CONTENTS

1. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of efficacy and safety data for Protocol SP-1011-002. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

This statistical analysis plan (SAP) is based on protocol version Amendment 2.1, dated 06 September 2017.

2. STUDY OBJECTIVES

2.1. PRIMARY OBJECTIVE

The primary objective of the study is to evaluate the efficacy (histological response) of APT 1011 in adults with eosinophilic esophagitis (EoE).

2.2. SECONDARY OBJECTIVES

The secondary objectives of the study are as follows:

- To define the dose-response of APT-1011
- To select a dose(s) of APT-1011 for Phase 3
- To evaluate the effect of APT-1011 on histology and endoscopic appearance
- To evaluate maintenance of efficacy and long-term safety of APT-1011
- To evaluate the population pharmacokinetics (PopPK) of APT-1011
- To evaluate the effect of APT-1011 on dysphagia episodes.

2.3. EXPLORATORY OBJECTIVES

[redacted]

3. STUDY DESIGN

3.1. GENERAL DESCRIPTION

FLUTE is a randomized, double-blind, placebo-controlled dose-ranging study of 4 total daily doses of APT 1011 versus placebo in 100 adult subjects (≥18 years of age) diagnosed with EoE.
The study will be conducted in several parts (Screening, 4-week single-blind placebo run-in and Baseline Symptom Assessment, and 2 treatment parts [Part 1 and Part 2]) with a Follow up Visit to occur 2 weeks after the final dose of study drug.

During the single-blind run-in/baseline symptom assessment, the subjects will receive placebo 30 minutes after breakfast and HS (at bedtime). At the end of the run-in period eligible subjects will be randomized to one of 4 doses of APT-1011 or placebo in a 1:1:1:1:1 ratio. Randomization will occur via an integrated Interactive Web Response System (IWRS), and will be stratified by the presence or absence of a history of or current esophageal stricture at Screening and history of a prior positive steroid response to any corticosteroid treatment previously received to treat the subject's EoE.

APT-1011 will be administered in 4 doses:

- 1.5 mg HS: Placebo 30 minutes after breakfast and 1.5 mg HS (at bedtime)
- 1.5 mg BID (1.5 mg 30 minutes after breakfast and at bedtime; total daily dose of 3 mg)
- 3 mg HS (Placebo 30 minutes after breakfast and 3.0 mg at bedtime)
- 3 mg BID (3.0 mg 30 minutes after breakfast and at bedtime; total daily dose of 6 mg)

[redacted]

Subjects randomized to placebo will receive matching placebo administered 30 minutes after breakfast and HS (at bedtime).
Efficacy (primary efficacy of histological response), safety, and PK of APT 1011 will be examined.

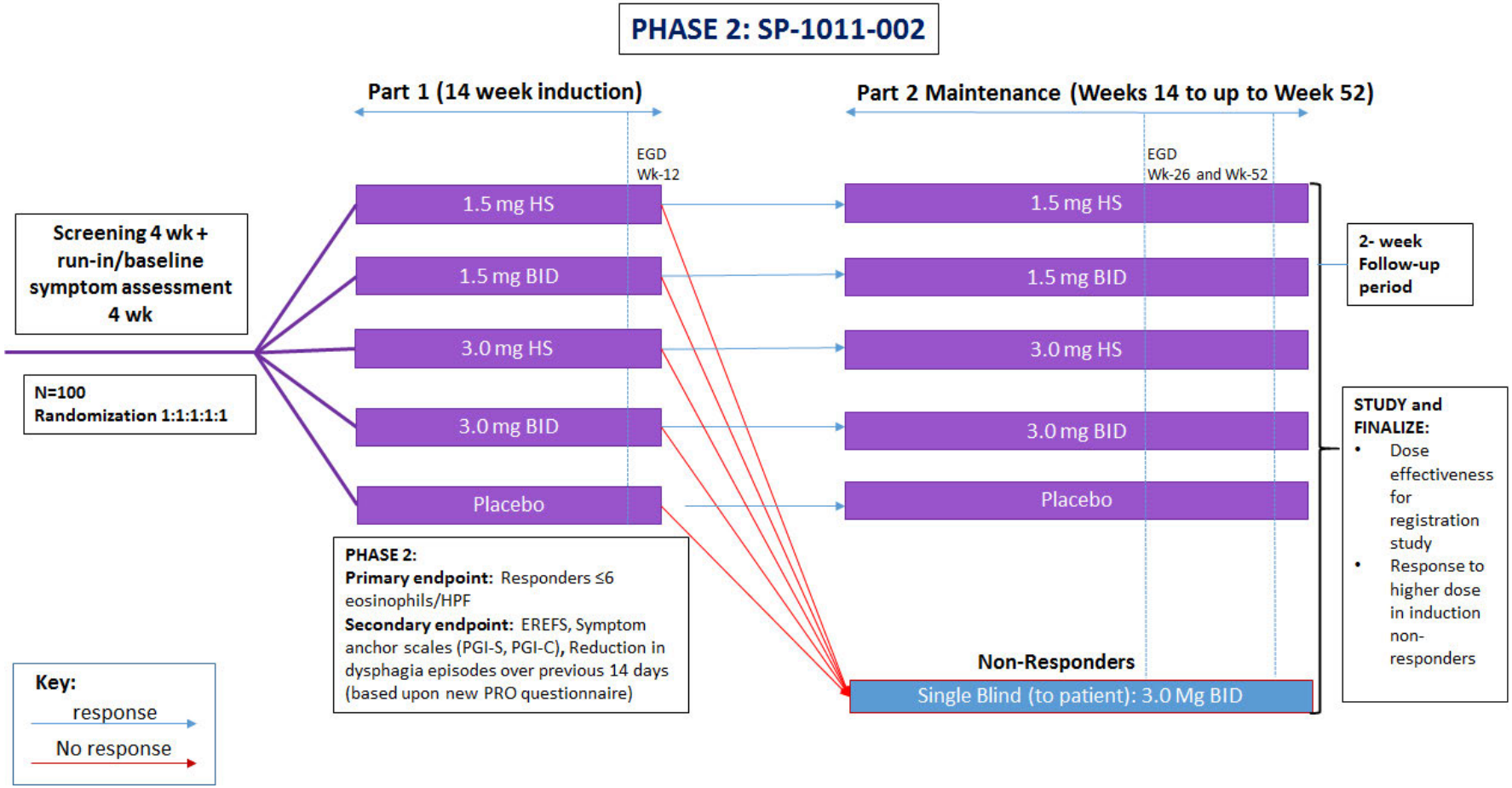
A histological assessment will be performed at Week 12, and non-responders to treatment will receive single-blind (to the subject) APT-1011 3.0 mg BID after Week 14, the end of Part 1 of the study. Part 2 of the study continues until Week 52. Subjects who are histological non-responders at Week 26 are to be withdrawn from the study at that time.

FLUTE is planned to be performed at approximately 60 active sites in North America (US and Canada) and Western Europe (Belgium, Germany, Switzerland and Spain).

Subjects who enter and complete FLUTE will be in the study for up to 62 weeks or until the last subject completes Week 28.

A study schematic is provided in Figure A.

Table A: Figure A Study Schematic



3.2. SCHEDULE OF EVENTS

Schedule of events can be found in Section 4.1.2 of the protocol.

3.3. CHANGES TO ANALYSIS FROM PROTOCOL

- Section 8.9 of the protocol states that "In subjects who change dosing groups, the TEAEs will be attributed to the previous dose, if they occur within 3 days of the change.". Given that subjects who change dose will always be changing from a lower or equivalent dose of study drug compared to the single-blind (to subject) APT-1011 3.0 mg BID dose which all subjects who change dose will receive, this 3 day rule will not be applied since it is not conservative and may result in more favorable safety results for the highest dose of study drug.
- Section 8.8.3 of the protocol states "As a secondary efficacy analysis, all 4 doses of active treatment will be pooled and compared to placebo with a one-sided test of two proportions." There was concern that this analysis could mask a significant effect at a particular dose, and so was excluded at this time. Pooling of effective doses may be considered for exploratory analyses at a later time if warranted.
- Section 8.9 of the protocol states that clinically significant changes in bone mineral density are to be summarized. However, this parameter was included in the protocol in error and is not captured in the study.
- Exploratory endpoints described in the protocol will be excluded from this SAP. After interim analysis they may be added to a separate SAP on a case-by-case basis as determined by the unblinded team. These outputs include:
- EoE sustained response (dysphagia): percentage of all subjects who met the dysphagia secondary endpoint at Week 12 and maintained a dysphagia-related response at Week 26 and Week 52
    - Evaluation of PK/PD (cortisol) and exposure-response (efficacy) relationships
    - Subject's assessment of symptoms compared with the previous visit
    - Change from baseline in EoE-QoL-A total score and subscores at Weeks 12, 26 and 52
    - Change from baseline in the number of dysphagia episodes for subjects who were classified as histologic non- responders at Week 12
- Traditionally, date of last dose and information regarding study drug compliance are recorded in the eCRF. However, due to electronic data collection design in the study, these were not collected in eCRF but instead were estimated using visit dates and an assumption of 100% compliance.
- Section 8.2 There will be an additional analysis population called Full Analysis Set (FAS). FAS will be the primary analysis population for efficacy. ITT will also be shown but not considerd the primary effiacy

population.

- Section 8.8.1 The gated analysis will be tested in the order of 3mg BID, 1.5mg BID, 3mg HS, 1.5mg HS and not follow what is in the protocol.

4. PLANNED ANALYSES

The following analyses will be performed for this study:

- Interim Analysis
- Final Analysis

The interim analysis will include analyses of the primary endpoint and selected secondary endpoints and will be performed when all randomized subjects have completed Week 12 of the study. Safety through Week 12 will also be assessed at this time.

The final analysis of the entire study will be performed when all randomized subjects have exited or completed the study.

4.1. DATA MONITORING COMMITTEE (DMC)

There will be no DMC for this study.

4.2. INTERIM ANALYSIS

- Analyses of the primary endpoint and selected secondary endpoints will be performed when all randomized subjects have completed Week 12 of the study. Safety through Week 12 will also be assessed at this time. All data (including all laboratory and non-CRF vendor data) collected in the project up to and including data until Visit 6/Week 12 will be considered for interim lock. Data for screen failed and early terminated subjects will also be included.
-

This will be the final analysis for primary and secondary endpoints for Weeks 12. There is no separate interim analysis plan, the SAP will be followed for the interim analysis as well as final analysis.

Safety analyses specified later in this document, are separated out into Part 1, Part 2 and overall (Parts 1 and 2 combined). The safety analyses planned for Part 1 will be performed for the interim analysis.

The list of statistical outputs and templates for these are provided with this SAP and outputs planned for the interim analysis are indicated as such within that document.

Once the programs have been produced by the IQVIA study team for blinded data, these programs will be sent to a separate, unblinded team, who will apply the randomization schedule and provide relevant people with a set of unblinded outputs. This will be detailed in an unblinding plan document, to be produced by IQVIA Biostatistics.

4.3. FINAL ANALYSIS

All final, planned analyses identified in this SAP will be performed by IQVIA Biostatistics following Sponsor Authorization of this Statistical Analysis Plan, Database Lock and Unblinding of Treatment.
Analyses relating to PK parameters will not be performed by IQVIA Biostatistics and will be addressed in a separate PK analysis plan.
Analyses related to the PROSE daily diary will be performed by a separate team and will not be covered in this SAP. The daily diary data will only be used by IQVIA Biostatistics for determining secondary and exploratory efficacy endpoints relating to dysphagia episodes, Global EoE, PGIC and PGIS.

5. ANALYSIS POPULATIONS

Agreement and authorization of subjects included/ excluded from each analysis population will be conducted prior to the unblinding of the study.

5.1. ALL SUBJECTS ENROLLED POPULATION [ENR]

The all subjects enrolled (ENR) population will contain all subjects who signed an informed consent form.

5.2. INTENT-TO-TREAT ANALYSIS POPULATION [ITT]

The intent-to-treat (ITT) analysis population will contain all subjects in the ENR population who were randomized and subjects will be classified according to randomized treatment.

5.3. FULL ANALYSIS SET POPULATION [FAS]

The full analysis set (FAS) analysis population will contain all subjects in the ITT population who have not met any of the following criteria:

1. Subjects who did not receive any study drug
2. Subjects given wrong drug
3. Subjects mis-randomized

5.4. PER PROTOCOL ANALYSIS POPULATION [PP]

The per-protocol (PP) analysis population will contain all subjects in the ITT analysis set who did not experience any reason for exclusion. It will be used for sensitivity analysis of the primary efficacy parameter.

Potential reasons for exclusion include:

- Entering study even though subject did not satisfy the inclusion/exclusion criteria
    - Inclusion/exclusion criteria violations will be as reported on the eCRF. IQVIA Biostatistics will not perform programmed checks of inclusion/exclusion criteria.
- Developing withdrawal criteria during the study and not withdrawing
- Receiving wrong treatment or incorrect dose
- Receiving a prohibited concomitant treatment (protocol sections 5.8.1/5.8.2) which may affect the efficacy of study drug
- Out of window Week 12 study visit/EGD, as defined in section 4.1.2 of the protocol.
- Non-compliance to study drug prior to the Week 12 EGD in the opinion of the investigator.
    - At each study visit the investigator or designee are to report non-compliance to study drug. Noncompliance is defined (section 5.10 of the protocol) as taking less than 80% or more than 120% of study drug during any evaluation period (visit to visit).

Protocol deviations will be reviewed by the Sponsor and IQVIA monthly. Decisions on whether deviations are to be considered major or minor will be documented at these meetings. Prior to the interim analysis, a list of major deviations will be provided, and decisions on subjects to be excluded from the PP analysis set will be finalized at an

analysis set review meeting and signed off by a representative of the Sponsor.

5.5. SAFETY ANALYSIS POPULATION [SAF]

The safety analysis set (SAF) will contain all subjects who receive at least one dose of study drug and subjects will be classified according to treatment received.
If there is any doubt whether a subject was treated or not, they will be assumed treated for the purposes of analysis.

See section 17 for the handling of safety data for Week 12 histological non-responders who switch to single-blind (to subject) APT-1011 3 mg BID at the end of Part 1 of the study.

5.6. SPARSE PK SUBGROUP

The Sparse PK Subgroup includes all subjects who have ≥1 quantifiable PK sample collected for sparse PK evaluations.
As stated in section 4.3 analyses relating to PK parameters will not be performed by IQVIA Biostatistics and do not form part of this SAP.

6. GENERAL CONSIDERATIONS

6.1. REFERENCE START DATE AND STUDY DAY

Study Day will be calculated from the reference start date, and will be used to show start/ stop day of assessments and events.

Reference start date is defined as the day of the first dose of study drug, (Day 1 is the day of the first dose of study drug), and Study Day will appear in every listing where an assessment date or event date appears. If first dose of study drug is missing, night of randomization will be assumed as Day 1.

For safety analyses of Part 2 of the study, histological non-responders at Week 12 who receive single-blind (to subject) APT-1011 3.0 mg BID during Part 2) will have an additional Study Day for Part 2 based on a reference start date which is day the first dose of study drug in Part 2 (i.e. Week 12 visit + 1 day).

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

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

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

  Study Day = (date of event – reference date).

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

6.2. BASELINE

Unless otherwise specified, baseline is defined as the last non-missing measurement taken prior to reference start date (including unscheduled assessments). In the case where the last non-missing measurement and the reference start date coincide, that measurement will be considered pre-baseline, but Adverse Events (AEs) and medications commencing on the reference start date will be considered post-baseline.

If any subjects are randomized but do not take study drug, their baseline as defined above will be defined relative to the date of randomization.
Note that for data other than adverse events (see section 17.1) baseline is always relative to the first dose of study drug in Part 1.

6.3. DERIVED TIMEPOINTS

There will be no derived timepoints in this study. EOT visits will be recorded separately and will not be re-mapped.

6.4. RETESTS, UNSCHEDULED VISITS AND EARLY TERMINATION DATA

In general, for by-visit summaries, data recorded at the nominal visit will be presented. Unscheduled measurements will not be included in by-visit summaries, but will contribute to any summaries of clinically significant or markedly abnormal safety data ‘at any time post-baseline’.

In the case of a retest (same visit number assigned), the earliest available measurement for that visit will be used for

by-visit summaries.

Listings will include scheduled, unscheduled, retest and EOT visit data.

6.5. WINDOWING CONVENTIONS

No windowing or mapping of visits will be conducted. Visits will be included as they are recorded by the site, regardless of study day.

6.6. STATISTICAL TESTS

The default significant level will be 5%; confidence intervals will be 90% and all tests will be one-sided.

6.7. COMMON CALCULATIONS

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

- Test Value at Visit X – Baseline Value

For certain parameters percentage change from baseline will be calculated as

- 100 x (Test Value at Visit X – Baseline Value)/Baseline Value

6.8. SOFTWARE VERSION

All analyses will be conducted using SAS version 9.4 or higher.

6.9. DATA SOURCES

All data will be retrieved from the eCRF with the following exceptions:

- Number of dysphagia episodes, PGIS, PGIC and Global EoE will be retrieved from the diary
- EoE histological response and EGD will be received from Inform Diagnostics
- Laboratory data will be received from Q2

- Randomization data will be received from Cenduit (after unblinding)

7. STATISTICAL CONSIDERATIONS

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

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

- History of or current presence of esophageal stricture (yes/no)
- Prior positive steroid response to any corticosteroid treatment previously received to treat EoE (yes/no)
- Geographic region (North America/Western Europe – see section 7.2)
- History of asthma/allergy
- EREFs score
- PPI status

7.2. MULTICENTER STUDIES

This study will be conducted by multiple investigators at multiple centers in North America and Western Europe. Randomization to dosing groups will occur in a double-blind manner using an integrated IWRS, and will be stratified by the presence or absence of a history of or current esophageal stricture at Screening and history of a prior positive steroid response to any corticosteroid treatment previously received to treat the subject's EoE captured with demography.

Randomization to dosing groups is not stratified by country/ center.
When specified, statistical analysis will be adjusted for geographic region or geographical region will be used as a subgroup. Geographic region will be categorized as follows:

Geographic Region	**Country**
North America	United States, Canada
Western Europe	Belgium, Germany, Switzerland and Spain

7.3. MISSING DATA

Missing efficacy data for the primary efficacy analysis will be imputed as described in sections 16.1.2 and 16.1.4 of this SAP.

Missing components of certain efficacy variables will be imputed prior to analysis/summarizing. This is described in the relevant secondary and exploratory endpoint analysis sections.

7.4. MULTIPLE COMPARISONS/ MULTIPLICITY

The primary analysis of the primary efficacy variable will be performed using a gatekeeping strategy to preserve Type I error. This strategy is described in section 16.1.3.

The analysis of the secondary efficacy parameter of number of dysphagia episodes will use a Holm-Step down procedure (Holm, 1979). This is described in section 16.2.3.4.

All other statistical testing performed is considered descriptive only and no adjustments will be performed for multiplicity.

7.5. EXAMINATION OF SUBGROUPS

Subgroup analyses will be conducted as stated in section 16.1.5. It should be noted that the study was not designed to detect treatment differences with high statistical power within subgroups.
The following subgroups will be assessed:

- History of or current presence of esophageal stricture:
    - Yes
    - No
- Prior positive steroid response to any corticosteroid treatment previously received to treat EoE:
    - Yes

- No

- Geographic region:
  - North America
  - Western Europe

- Gender:
  - Female
  - Male

- Age (years):
  - 18 - < 30 years
  - 30 - < 45 years
  - 45 - < 65 years
  - > = 65 years

- Race in 3 categories:
  - Black/African American or Afro-Carribbean
  - White
  - All other races combined (American Indian/Alaska Native, Asian, Native Hawaiian/Other Pacific Islanders)

- Ethnicity:
  - Hispanic or Latino
  - Not Hispanic or Latino

- Proton Pump Inhibitor (PPI) status:
  - Continuing into the study
  - Not continuing into the study

8. OUTPUT PRESENTATIONS

Appendix 1 shows conventions for presentation of data in outputs.

The templates provided with this SAP describe the presentations for this study and the format and content of the summary tables, figures, and listings to be provided by IQVIA Biostatistics.

9. DISPOSITION AND WITHDRAWALS

All subjects who provided informed consent will be accounted for in this study.
Subject disposition and study discontinuations, including reasons for study discontinuation will be presented for the ENR set for the placebo run-in period prior to randomization.

Subject disposition and study discontinuations, including reasons for study discontinuation during Parts 1 and Part 2 will be presented for the ITT and FAS analysis sets. The number of histologic non-responders at Week 12 who received single-blind (to subject) APT-1011 3.0 mg BID in Part 2 will be presented.
For Part 2, disposition information will also be presented separately for histologic non-responders at Week 12 who received to single-blind (to subject) APT-1011 3.0 mg BID in Part 2.

Analysis set disposition and reasons for exclusion from the PP analysis set will be presented for all subjects in the ENR set.

Protocol deviations will be listed.

10. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic data and other baseline characteristics will be presented for each of ENR set, the ITT analysis set, the FAS analysis set, and the SAF.
No statistical testing will be carried out for demographic or other baseline characteristics.
The following demographic and other baseline characteristics will be reported for this study:

- Sex (Male/Female)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Other)
- Ethnicity (Hispanic or Latino/Not Hispanic or Latino)
- Age (years) - calculated relative to date of informed consent
- Geographic region (North America/Western Europe)
- Weight (kg)

- Height (cm)
- Body mass index (BMI) (kg/m^2)
- Smoking status (Never/Former/Current)
- Global EoE symptom score
- History of esophageal stricture(s) (Yes/No)
- Current esophageal stricture(s) based on the study esophagogastroduodenoscopy (EGD) (Yes/No)
- History of a positive steroid response to EoE treatment (Yes/No)
- PPI status (Continuing Yes/No)

10.1. DERIVATIONS

- BMI (kg/ m^2) = weight (kg)/ height (m)^2

11. SURGICAL AND MEDICAL HISTORY

Surgical and medical history information will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary, Version 21.0 and presented for the SAF.

- Surgical procedures captured on the *Medical and Surgical History* page of the eCRF will be presented by SOC (System Organ Class) and PT (Preferred Term).
- Medical history conditions are defined as those conditions which stop prior to or at Screening. Medical history conditions captured on the *Medical and Surgical History* page of the eCRF will be presented by SOC and PT.

12. CONCOMITANT ILLNESSES

Concomitant Illnesses will be coded using the MedDRA coding dictionary, Version 22.0 and presented for the SAF.

- Concomitant Illnesses are medical conditions (other than EoE) which started prior to or at Screening and are ongoing at the date of Screening. Concomitant illnesses captured on the *Medical and Surgical History* page of the eCRF will be presented by SOC and PT.
- Concomitant illnesses will be included in the medical history summary, and indicated as concomitant in the listings.

13. MEDICATIONS

Medications will be coded using the WHO Drug Dictionary (WHODD), 01 September 2017 and presented by ATC Level 3 and Preferred Term for the SAF. A modified Preferred Term has been provided by Adare and will be presented in the summary tables.

See Appendix 2 for handling of partial dates for medications, in the case where it is not possible to define a medication as prior, concomitant, or post treatment, the medication will be classified by the worst case; i.e. concomitant.

- 'Prior' medications are medications which started and stopped prior to the first dose of study drug.
- 'Concomitant' medications are medications which:
    - started prior to, on or after the first dose of study drug
    - AND ended on or after the date of first dose of study drug or were ongoing at the end of the study.
- 'Post' medications are medications which started after the last dose of study drug.

14. STUDY DRUG EXPOSURE

Exposure to study drug in weeks will be presented for the SAF.
It will be derived separately for each of Part 1, Part 2 and for Parts 1 and Parts 2 combined.

- For Part 1 all subjects will be included in their double-blind dosing groups.
- For Part 2, histologic responders at Week 12 who continue into Part 2 will be included in their double-blind dosing groups.
- For Part 2, histologic non-responders at Week 12 who continue into Part 2 will be included in a single-blind APT-1011 3.0 mg BID group. Additionally, a summary will be provided for all APT-1011 3.0 mg BID during Part 2, both double-blind and single-blind.
- For Parts 1 and 2 combined, histologic non-responders at Week 12 will be included in their double-blind dosing group (counting just their Part 1 exposure) and also in the single-blind APT-1011 3.0 mg BID group (counting just their Part 2 exposure). A summary will be provided for all APT-1011 3.0 mg BID during Part 2, both double-blind and single-blind.

The date of first study drug administration in Part 1 will be taken from the eCRF " Sparse Pharmacokinetic Sampling" form. The date of last study drug will be estimated using study visits. For subjects who complete the

Week 12 visit, the last dose of study drug for part 1 will be assumed to be on the morning of the date of that visit. For subjects who complete the Week 52 visit, the last dose of study drug for part 2 will be assumed to be the evening before this visit. For subjects who discontinue the trial prior to either of these dates, date of last dose will assumed to be their early termination visit.

Compliance will be assumed to be 100%, interruptions and dose changes are not taken into account for duration of exposure.

Study drug exposure time will be divided into 5 categories, such that an equal number of subjects falls into each catgory (i.e. 20th, 40th, 60th, and 80th percentiles of exposure time will be used as cut-offs).

14.1. DERIVATIONS

Duration of exposure (weeks) for Part 1=
(date of last study drug administration in Part 1 – date of first study drug administration in Part 1 + 1) / 7.
Duration of exposure (weeks) for Part 2=
(date of last study drug administration in Part 2 – date of first study drug administration in Part 2 + 1) / 7.
Duration of exposure (weeks) for Parts 1 and 2 combined, for histological responders at Week 12 =
(date of last study drug administration in the study – date of first study drug administration in Part 1 + 1) / 7.

Total exposure to study drug for Part 1 =
(Duration of exposure (days) * total daily dose) – evening dose.
Duration of exposure for Part 2 =
(Duration of exposure (days) * total daily dose)

15. STUDY DRUG COMPLIANCE

Compliance to study drug will be summarized for the ITT and FAS analysis sets.

Detailed study drug accountability data will not be available to IQVIA Biostatistics.

Compliance is to be assessed by the investigator and their assessment of compliance at each visit reported as either

compliant or non-compliant. The protocol defines non-compliance as ‘taking less than 80% or more than 120% of study drug during any evaluation period (visit to visit)’.

Summaries of proportions of compliant/non-compliant subjects at each study visit will be provided. These summaries will not include data for histologic non-responders at Week 12 after they receive single-blind (to subject) APT-1011 3.0 mg BID at the start of Part 2. i.e. only information while subjects are taking their randomized study drug will be included.

16. EFFICACY OUTCOMES

16.1. PRIMARY EFFICACY

16.1.1. PRIMARY EFFICACY VARIABLE & DERIVATION

The primary efficacy variable is the following histological endpoint:

- Percentage of subjects with ≤6 peak eosinophils/HPF after assessing at least 5-6 biopsies from the proximal and distal esophagus (~3 each) where the HPF area is 235 square microns (40 magnification lens with a 22 mm ocular).

The assessment or response is performed by the pathologist and will not be derived by IQVIA Biostatistics.

16.1.2. MISSING DATA METHODS FOR PRIMARY EFFICACY VARIABLE

For the primary analysis of the primary efficacy variable, any subject who does not meet the definition of responder, will be classified as a non-responder. This includes subjects who withdraw from the study prior to the week 12 histological assessment.
The primary analysis of the primary efficacy variable is via a stratified Cochran-Mantel-Haenzel (CMH) test, as described in section 16.1.3.

A sensitivity analysis, described in section 16.1.4 will be performed with missing response data imputed using multiple imputation, assuming the data are missing at random.
Multiple imputation will be used to provide estimates of the common odds ratio, its 90% confidence interval and associated 1-sided p-value.

The dataset containing subject, dosing group, week 12 histologic response, the two randomization strata, history of asthma/allergy, baseline EREF score, and PPI status will be used as the basis for multiple imputations. Histologic response will be set to missing for those subjects who did not have a Week 12 result for peak eosinophils/HPF (and for whom response status was set to non-responder in the primary efficacy analysis).

The methodology which follows in this section are taken from a 2013 paper by Ratitch, Lipkovich and O'Kelly.
The imputation method will be based on logistic regression using the MONOTONE statement of PROC MI in SAS

as shown below.

The option NIMPUTE in the PROC MI statement will be used to specify the number of imputed datasets to be generated. The output dataset will contain 500 copies of the original dataset, with the observed values being the same across all datasets, and with imputed values varying from one dataset to another. These multiple copies will be identified by a new variable, _Imputation_, added to the output dataset by PROC MI.
The SAS Code to perform the imputations is provided below. The SEED and NIMPUTE value are explicitly stated here to enable results to be replicated and quality assured.

```
PROC MI DATA=datain OUT=datain mi SEED=1201913 NIMPUTE=500;
VAR dosegrp strata1 strata2 asthma EREF PPI response;
CLASS dosegrp strata1 strata2 response;
MONOTONE LOGISTIC;
RUN;
```

where dosegrp has 5 levels corresponding to the 5 dose groups (including placebo), strata1 and strata2 are the binary randomization strata, asthma is a binary variable for history of asthma/allergy (yes/no), and EREF is continuous variable for EREF score at baseline, PPI is a binary variable for PPI status (continuing into study vs not) and response is the binary response variable.

After imputation is performed, the next step will be analyzing each of the imputed datasets. This will be done using the same method (stratified CMH test) that was used for the primary efficacy analysis.

The SAS Code provided below is for one of the pairwise tests, assuming that dosegrp = 1 represents the APT-1011 3.0 mg BID group and dosegrp = 5 represents the placebo group. All of the analyses of imputed datasets described below and the combining of relevant statistics to provide overall imputed estimates of odds ratio, 90% confidence interval and 1-sided p-values would need to be performed separately for each of the four APT-1011 dosing groups and placebo.

```
*** Obtain Mantel-Haenzel estimate of the common odds ratio for APT-1011 3mg
BID vs Placebo adjusted for
randomization strata ***;
PROC FREQ DATA=datain mi(where=(dose in (1,5)));
TABLES strata1*strata2*dosegrp*response/CMH;
ODS OUTPUT COMMONRELRISKS=comrrout;
BY  Imputation_;
RUN;
```

The estimates of odds ratios follow a log-normal distribution and a log transformation will be applied to normalize

these estimates in order to be able to apply Rubin's (Rubin, 1987) combination rules. These combination rules take as input estimates of a statistic obtained from multiple imputed datasets as well as standard errors of these estimates, and produce an overall pooled estimate and an overall standard error and confidence interval.

The SAS code provided below shows a log transformation applied to the estimates of the odds ratio for the treatment effect. Standard error of the transformed estimate is obtained from the log-transformed lower and upper confidence limits for the odds ratio estimate. Then the dataset containing the transformed estimates and their standard errors is passed to PROC MIANALYZE. The MODELEFFECTS statement contains a name of the variable that represents an estimate of the statistic to be combined, and the STDERR statement contains the name of the variable that represents standard errors of that estimate. The combined results are captured in the dataset PARAMETERESTIMATES. The combined estimate of the odds ratio can then be back-transformed to its original log scale as shown in the last data step of SAS Code, which also computes 90% confidence limits on the log scale using the combined estimate of the standard error for the odds ratio.

```
*** Log-transform odds ratio estimates
and obtain standard error from confidence intervals ***;
DATA ormh t;
SET comrrout(WHERE=(StudyType="Case-Control"));
log_or_mh_value=log(VALUE);
log or_mh_se=(log(UPPERCL)-log(LOWERCL))/(2*1.96);
RUN;

*** Combine transformed estimates;
PROC MIANALYZE DATA=ormh t;
ODS OUTPUT PARAMETERESTIMATES=mian_ormh_t;
MODELEFFECTS log or mh_value;
STDERR log_or_mh_se;
RUN;

*** Back-transform combined values;
DATA mian_mhodds_bt;
SET mian lgsodds t;
Estimate back = EXP(ESTIMATE); *Pooled odds ratio;
LCL_back=Estimate_back*EXP(-1.645*STDERR); *Pooled lower limit;
UCL back=Estimate_back*EXP(+1.645*STDERR); *Pooled upper limit;
RUN;
```

The CMH general association statistic test, under the null hypothesis of no association between APT-1011 dose and placebo, having controlled for the randomization strata has an asymptotic chi-square distribution with 1 degree of freedom. The chi-square distribution with 1 degree of freedom is highly skewed and obtaining a combined result of the CMH test from multiply-imputed data requires a transformation that would normalize the CMH statistic.

The Wilson-Hilferty transformation (Wilson & Hilferty, 1931; Goria, 1992) can be used for this purpose:

$$wh\_cmh^{(m)} = \sqrt[3]{cmh^{(m)}/df} \quad (1)$$

where $cmh^{(m)}$ is the CMH statistic computed from the mth imputed dataset, df is the number of degrees of freedom associated with the CMH statistic, and $wh\_cmh^{(m)}$ is the transformed value. The transformed statistic is approximately normally distributed with mean $1 - 2/(9 \times df)$ and variance $2/(9 \times df)$ under the null hypothesis.

The transformed statistic in (1) can be standardized to obtain a variable that is normally distributed with mean 0 and variance 1:

$$st_{wh\_cmh^{(m)}} = \frac{\sqrt[3]{cmh^{(m)}/df} - \left(1 - \frac{2}{9 \times df}\right)}{\sqrt{\frac{2}{9 \times df}}} \quad (2)$$

This transformed statistic can be passed to PROC MIANALYZE in SAS in order to perform a combined CMH test. The SAS code below contains an invocation of PROC FREQ to request the CMH test using the CMH option in the TABLES statement, with the results captured in the ODS output dataset CMH. A subsequent data step applies the Wilson-Hilferty transformation as described in equation (2) and then passes the transformed values to PROC MIANALYZE using the same syntax as for the odds ratio.

Finally, a p-value (probt_upper) for the combined CMH test can be obtained as the upper-tailed p-value from the normal test produced by PROC MIANALYZE on the transformed statistic. This is done in the last data step of the SAS Code below.

```
*** Perform CMH test;
PROC FREQ DATA=datain_mi(where=(dose in (1,5)));
TABLES strata1*strata2*dosegrp*response/CMH;
ODS OUTPUT CMH=cmh;
BY _Imputation_;
RUN;

*** Apply Wilson-Hilferty transformation to the CMH statistic and
standardize the resulting normal variable;
DATA cmh wh;
SET cmh(WHERE=(AltHypothesis="General Association"));
```

```
cmh value wh=((VALUE/DF)**(1/3) - (1-2/(9*DF)))/SQRT(2/(9*DF));
cmh_sterr_wh = 1.0;
RUN;

*** Combine results;

PROC MIANALYZE DATA=cmh wh;
ODS OUTPUT PARAMETERESTIMATES=mian_cmh_wh;
MODELEFFECTS cmh_value_wh;
STDERR cmh_sterr_wh;
RUN;
*** Compute one-sided p-value;
DATA mian cmh wh_p;
SET mian cmh wh;
IF tValue > 0 THEN Probt_upper = Probt/2;
ELSE Probt_upper = 1-Probt/2;
RUN;
```

16.1.3. PRIMARY ANALYSIS OF PRIMARY EFFICACY VARIABLE

The primary analysis of the primary efficacy variable (histological responder or non-responder) will be performed with a stratified CMH test using the two randomization strata: history of or current presence of esophageal stricture (yes/no) and prior positive steroid response to any corticosteroid treatment previously received to treat EoE (yes/no). There are 4 hypotheses corresponding to the 4 APT-1011 dosing groups, which will be tested using a gatekeeping strategy to preserve Type I error for each analysis.

Let θj be the CMH common odds ratio comparing treatment vs. control for histologic response for dose j given the randomization strata; with j=1,2,3,4 corresponding to APT-1011 3 mg BID, 1.5 mg BID, 3 mg HS, 1.5 mg HS doses, respectively.

- Primary hypothesis #1

H_0: $\theta_1 \leq 1$

H_1: $\theta_1 > 1$

If the corresponding one-sided p-value from the CMH test of primary hypothesis #1, i.e., APT-1011 3 mg BID vs. placebo, is ≤0.05, the null hypothesis will be rejected, and subsequently the following hypothesis will be tested:

- Primary hypothesis #2

H_0: $\theta_2 \leq 1$

H_1: $\theta_2 > 1$

If the corresponding one-sided p-value from the CMH test of primary hypothesis #2, i.e., APT-1011 1.5 mg BID vs. placebo, is ≤0.05, the null hypothesis will be rejected, and subsequently the following hypothesis will be tested:

- Primary hypothesis #3

H_0: $\theta_3 \leq 1$
H_1: $\theta_3 > 1$
If the corresponding one-sided p-value from the CMH test of primary hypothesis #3, i.e., APT-1011 3 mg HS vs. placebo, is ≤0.05, the null hypothesis will be rejected, and subsequently the following hypothesis will be tested:

- Primary hypothesis #4

H_0: $\theta_4 \leq 1$
H_1: $\theta_4 > 1$
If the corresponding one-sided p-value from the CMH test of primary hypothesis #4, i.e., APT-1011 1.5 mg HS vs. placebo, is ≤0.05, the null hypothesis will be rejected.
Note the gate-keeping strategy only allows formal hypothesis testing of each dose if the higher doses all meet statistical significance. Given prior *studies* with histologic endpoints, there is a strong rationale suggesting that higher doses will have a greater benefit on the primary efficacy endpoint of histological response.

The common odds ratio, 90% confidence interval and 1-sided p-value will be from those produced using the Mantel-Haenszel method from the FREQ procedure in SAS.

The primary efficacy analysis will be performed for the FAS analysis set.

16.1.4. SENSITIVITY ANALYSIS OF PRIMARY EFFICACY VARIABLE(S)

- Sensitivity to analysis set
    - The primary analyses of the primary efficacy variable will be repeated on the PP analysis set.
- Sensitivity to adjustment for covariates
    - The primary efficacy variable will be analyzed using logistic regression. Each logistic regression model will contain the subjects randomized to each of the four APT-1011 dosing groups and the patients randomized to placebo. Histological response will be the response variable, and the covariates and factors in the model will be all of those listed in section 7.1. Other than the two

randomization strata, co-variates will only be added using stepwise selection, and will be added at 10% significance level and removed at 15% significance level. These logistic regression analyses will be performed on the FAS analysis set.

- Sensitivity to missing data assumptions
    - The primary efficacy variable will be analyzed using multiple imputation methods, assuming the data are missing at random (MAR). See section 16.1.2.

16.1.5. SECONDARY ANALYSES OF THE PRIMARY EFFICACY VARIABLE

- Bayesian hierarchical modeling will be used for both the primary efficacy endpoint and reduction in number of dysphagia episodes, in which dose-response modeling is used to estimate the difference between each APT-1011 dosing group and placebo.
- Terms in the Bayesian models will include dose (3 mg, 1.5 mg, or 0 mg) , frequency (BID or HS, with placebo subjects included as BID), and the two randomization strata. Posterior estimates for each of the model coefficients will be generated through Monte Carlo Markov Chains (MCMC) using the Metropolis random walk sampling algorithm. Initially the first 2000 simulations will be discarded for burn-in and the subsequent 10000 samples retained for posterior estimation. Trace iteration plots and autocorrelation plots of the simulated samples will be examined with standard diagnostic tests to establish convergence of the posterior samples. If sufficient convergence of all parameters has not been achieved the burn-in period and number of simulations will be revised and different sampling algorithms will be considered.
- The proportion of responders at each dose level and frequency will be estimated from the respective posterior distributions. Estimates of the difference in the proportion of responders between each active dose level and placebo will be calculated using the distributions of the differences between posterior estimates along with 90% credible intervals of these estimates. The probability that the difference is greater than 0 will also be displayed. In addition the posterior distributions of each difference estimate will be produced along with the estimated dose-response curve. Estimates of the dosing frequency and stratification effects will also be displayed for all dose groups.
- A logistic regression model using both dose and frequency and dose-frequency interaction will be used to analyze the primary efficacy variable. Separate terms will be included in the model for dose (3 mg, 1.5 mg or 0 mg (placebo) frequency (BID or HS, with placebo subjects included as BID) and dose-frequency interaction.

Two models will be used. One containing just dose, frequency and dose-frequency interaction. The other will

additionally contain randomization strata.

- Odds ratios will be provided per 1.5 mg increase in dose and per 1 unit increase in daily dose frequency. 3 mg doses will be coded as 2, 1.5 mg doses will be coded as 1, and placebo will be coded as 0. For frequency, BID will be coded as 2, and HS will be coded as 1. The variables will be analyzed in continuous form, but re-coding in this way will allow for provision of the odds ratios per 1.5 mg increase.
- If the dose-frequency interaction term is significant at the 10% level, separate logistic regression analyses models will be fitted for each dose level separately, with dose frequency in the model (and repeated with the randomization strata also in the model). If the interaction term is not significant, a model without the interaction terms will be used to estimate the effects of dose and frequency

- The primary efficacy variable will be summarized for each subgroup defined in section 7.5, by each of the five randomized dosing groups.
  - For the dose with maximum utility (i.e. largest effect size), logistic regression will be used to estimate Odds ratio of APT-1011 vs Placebo within each subgroup
  - For other doses, no statistical testing will be performed for results within each subgroup.

All these secondary analyses of the primary efficacy variable will be performed on the FAS analysis set.

16.2. SECONDARY EFFICACY

16.2.1. SECONDARY EFFICACY VARIABLES & DERIVATIONS

16.2.1.1. EoE sustained response: maintenance of ≤6 peak eosinophils/HPF at Weeks 26 and 52 for histologic responders at Week 12

For histologic responders at Week 12, maintenance of ≤6 peak eosinophils/HPF at each of Week 26 and Week 52 is defined as having ≤6 peak eosinophils/HPF (as per the primary efficacy variable in section 16.1.1), at each of Weeks 26 and 52 respectively.

Subjects who withdrew prior to each of Weeks 26 and 52 will be classified as non-responders at those visits.

16.2.1.2. Change from baseline EREFs Total Score at Week 12, Week 26, and Week 52

For total esophagus:

- each of edema and stricture will be classified as either Grade 0 (absent) or Grade 1 (present)
- rings will be classified as Grade 0 (none), Grade 1 (mild), Grade 2 (moderate), Grade 3 (severe)
- each of exudates and furrows will be classified as Grade 0 (none), Grade 1 (mild), Grade 2 (severe)

Crepe paper esophagus and narrow caliber esophagus will each be classified as Grade 0 (absent) or Grade 1 (present).

Esophageal erosions will be classified as Normal, and each of Grade A through to Grade D based on the Los Angeles Classification, but not separately for the proximal and distal esophagus.

A total EREFs score will be derived as follows:

Component	Grade	Score
Edema*	Absent Present	0 1
Rings*	Grade 0 (none) Grade 1 (mild) Grade 2 (moderate) Grade 3 (severe)	0 1 2 3
Exudates*	Grade 0 (none) Grade 1 (mild) Grade 2 (severe)	0 1 2
Furrows*	Grade 0 (none) Grade 1 (mild) Grade 2 (severe)	0 1 2
Stricture*	Absent Present	0 1

Crepe paper esophagus	None Present	0 1
Narrow caliber esophagus	None Present	0 1
Esophageal erosions (Los Angeles Classification)	Normal Grade A Grade B Grade C Grade D	0 1 2 3 4
Total score		Sum of the above (ranging from 0 to 15)

- * For this component the grade/score will be based on the WORST result.
- The change from baseline in the EREFs total score will be summarized as continuous variables, and analysis of covariance (ANCOVA) will be used to compare the mean change from baseline in each of the APT-1011 dosing groups to placebo, at the Week 12 visit, with the estimated mean difference, 90% confidence interval of the mean difference and 1-sided p-values provided. This statistical testing at Week 12 is considered descriptive only and no adjustments will be made for multiple comparisons.
- The ANCOVA models will include the factors described in section 7.1.
-

16.2.1.3. Peak eosinophils/HPF number <1 and <15 at Week 12, Week 26, and Week 52

Subjects who had withdrawn prior to each visit will not have results imputed, and percentages will be based on the number of subjects with EGD results assessed at the visit.

Subjects with a result of <1/HPF will also be counted as having a result of <15/HPF.

16.2.1.4. Change from baseline Global EoE Symptom Score at each post-baseline visit

The global EoE symptom score, ranging from 0 to 10 will be provided directly to IQVIA Biostatistics with no derivation required. This will be retrieved from the diary.

16.2.1.5. Change from baseline in the number of dysphagia episodes at Week 12, Week 26 and Week 52

The number of dysphagia episodes for a subject will be determined from the 14 day evening diary completed prior

to each of Week 12, Week 26 and Week 52. The number of episodes for each day will have 2 components a real time episode component along with the CRF entry of the number of dysphadia free days in the last 24 hours. First a day must be identified as a valid day. A day is considered valid if the subject awnsers yes to the CRF question "In the last 24 hours, did you have any difficulty with food or pills going down."

With a valid day from the evening diary the subject will add the real time episode component which comes from the swallowing difficulty. This component is the sum for each day of the CRF question "Did you just experience difficulty with food or pills going down?" The days will need to be readjusted if the time that they awnser the swallowing question and the time that they do the evening diary is more than 30 minutes difference. If the difference from episode time to diary time is greater than 30 minutes then this episode will be considered to have happened on the previous day.

Total dysphagia episodes will be the sum of the real time events for the previous 14 days of the subjects scheduled visit date.

A dysphagia free day will be classified if there are no real time events and no recalled episodes. Total dysphagia free days will be calculated as the sum of the previous 14 days of the subject scheduled visit date.

16.2.1.6. Change from baseline in 7-day EEsAI total score at Weeks 12, 26 and 52

The total score for the Eosinophilic Esophagitis Activity Index (EEsAI) will be derived as follows from the five separate components, described in section 16.2.1.7.

Item	Categorized Result	Score
Frequency of trouble swallowing	Never	0
	1-3 times/week	15
	4-6 times/week	27
	Daily	31
Duration of trouble swallowing	≤5 minutes	0
	>5 minutes	6
Pain when swallowing	No	0
	Yes	15
VDQ score	0	0

Item	Categorized Result	Score
	>0 – 2.5	12
	>2.5 – 5.0	19
	>5.0 – 7.5	21
	7.5 – 10.0	23
AMS score	0 – 5.0	0
	>5.0 – 7.5	9
	>7.5 – 10.0	25
Total score		100

16.2.1.7. Change from baseline in 7-day EEsAI subscores at Weeks 12, 26 and 52

There are five subscores of the EEsAI:

- Frequency of trouble swallowing
- Duration of trouble swallowing
- Pain when swallowing
- Visual Dysphagia Question (VDQ) score
- Avoidance, Modification and Slow Eating (AMS) score

Subjects were asked whether they had, in the past 7 days, eaten the following types of food:

- Solid meat (steak, chicken, turkey, lamb)
- Soft foods (pudding, jelly, apple sauce)
- Dry rice (grains don't stick) or sticky Asian rice (without sauce)
- Ground meat (hamburger, meat loaf)
- Fresh white untoasted bread or similar foods (donut, muffin, cake)
- Grits, porridge (oatmeal), rice pudding
- Raw, fibrous foods, not grated (apple, carrot, celery)
- French fries without sauce or ketchup

Subjects were shown pictures of the 8 types of food and were asked to assess how difficult it would be to swallow

each of them: The degree of perceived difficulties when eating a given food consistency will be scored as 0 for No difficulties, 1 for mild difficulties, 2 for moderate difficulties and 3 for severe difficulties. These scores for each food consistency were summed in the numerator of the score and divided by the maximum sum of grades that could be attained for each subject, which depends on the number of food consistencies consumed by a subject in a given recall period.

The VDQ score is derived as:

$$\text{VDQ} = \frac{(\text{N1 x 1}) + (\text{N2 x 2}) + (\text{N3 x 3})}{\text{D x 3}} \text{ x } 10$$

Where:

1. N1 = number of food consistencies graded with 'Mild difficulties'
2. N2 = number of food consistencies graded with 'Moderate difficulties'
3. N3 = number of food consistencies graded with 'Severe difficulties'
4. D = number of relevant food consistencies for that subject

For the AMS score, answers to three items exploiting the pattern of behavioral adaptation were scored for each food consistency consumed by the subject. If patients recorded no behavioral changes, a score of 0 was assigned; when reporting eating slower than others, a score of 1 was assigned; when reporting the modification of certain food consistencies, a score of 2 was assigned; when reporting both eating slower than others and modifying certain food consistencies, a score of 3 was assigned; if the subject completely avoided one or several food consistencies due to EoE symptoms, a score of 5 was assigned. Scores for all consumed food consistencies were summed up in the numerator and divided by the maximum sum of scores that could be attained by a given subject.

The AMS score is derived as:

$$\text{AMS} = \frac{(\text{N1 x 1}) + (\text{N2 x 2}) + (\text{N3 x 3}) + (\text{N4 x 5})}{\text{D x 5}} \text{ x } 10$$

Where:

1. N1 = number of food consistencies with 'Yes' to 'Eating slowly' only

2. N2 = number of food consistencies with 'Yes' to 'Modification' only
3. N3 = number of food consistencies with 'Yes' to both 'Eating slowly' and 'Modification'
4. N4 = number of food consistencies with 'Yes' to 'Avoidance' only
5. D = number of relevant food consistencies for that subject.

16.2.1.8. Subjects with 7-day EEsAI total score <20 at Weeks 12, 26 and 52

EEsAI total score will be derived as described in section 16.2.1.6.

16.2.1.9. Change from baseline PGIS at each post-baseline visit

The Patient Global Impression of Severity (PGIS) comprises two questions:

- severity of your EoE symptoms over the past 7 days
- severity of your difficulty with food or pills going down over the past 7 days
- For each question, subjects can answer:
- None
- Mild
- Moderate
- Severe
- Very Severe.
- This information will be retrieved from the daily diary;

16.2.1.10. PGIC at each post-baseline visit 52

The Patient Global Impression of Change (PGIC) comprises the same two questions as the PGIS (see section 16.2.1.9):

- severity of your EoE symptoms
- severity of your difficulty with food or pills going down

with subjects assessing themselves relative to before starting study treatment.

PGIC responses at each post-baseline visit will be one of

- Much worse
- Moderately worse
- A little worse
- Stayed the same

- A little improved
- Moderately improved
    - Much improved.
    - This information will be retrieved from the daily diary;

16.2.1.11. Assessment of treatment failure and relapse, including

- Percentage of histologic non-responders at Week 12, Week 26, and Week 52

Histologic non-responders are subjects who do not achieve ≤6 peak eosinophils/HPF after assessment of at least 5-6 biopsies from the proximal and distal esophagus (~3 each) where the HPF area is 235 square microns (40 magnification lens with a 22 mm ocular).
For each visit, subjects who have withdrawn from the study prior to the visit are classified as non-responders at that visit.

- Percentage of subjects requiring emergency endoscopic food dis-impaction by dose before Week 14, between Week 14 and Week 28, and between Week 28 and Week 52 – retrieved from EGD assessment before unscheduled visit
- Percentage of subjects requiring esophageal dilation by dosing group and part of the study – retrieved from EGD assessment before unscheduled visit

16.2.2. MISSING DATA METHODS FOR SECONDARY EFFICACY VARIABLES

Subjects who did not complete Week 26 histological assessment will be classified as Week 26 non-responders. Subjects who were non-responders at Week 26 are to be withdrawn from the study at Week 26. These subjects and any subjects who either terminated the study early or did not completed Week 52 histological assessment will be counted as non-responders at Week 52.

Visitwise missing data will not be imputed in summaries or analyses of exploratory variables.
Rules for imputing missing data for the derivation of other individual parameters prior to summarizing or analyzing the variable, are provided in the relevant sections **Error! Reference source not found.** through to **Error! Reference source not found.**.

16.2.3. ANALYSIS OF SECONDARY EFFICACY VARIABLES

Statistical tests to compare each APT-1011 dosing group with placebo will be performed for some secondary efficacy endpoints as described in the following sections, but the corresponding p-values will be considered as descriptive rather than inferential. The analyses and summaries of secondary efficacy variables will be performed for FAS analysis set.

16.2.3.1. Analysis of EoE sustained response: maintenance of ≤6 peak eosinophils/HPF at Weeks 26 and 52 for histologic responders at Week 12

The frequency and percentage of histologic responders at Week 12 who maintained response at each of Week 26 and Week 52 will be provided for each randomized dosing group. Only histologic responders at Week 12 will be included in the summary.

No statistical testing will be performed.

16.2.3.2. Analysis of change from baseline PGIS and change from baseline in each of the EEsAI items trouble swallowing, duration of trouble swallowing and pain when swallowing

For each of the two PGIS questions (see section 16.2.1.9) and for the EEsAI items trouble swallowing frequency, duration of trouble swallowing and pain when swallowing, shift tables will be provided for the results at baseline and at each post-baseline visit, for each randomized dosing group. Data will not be not included for visits after a histologic non-responder at Week 12 had received single-blind (to subject) APT-1011 3 mg BID in Part 2.

No statistical testing will be performed.

16.2.3.3. Analysis of percentage of subjects with a peak eosinophils/HPF number <1 and <15 and percentage of subjects with 7-day EEsAI total score <20

The number and percentage of subjects in each of these categories will be presented by visit for each of the randomized dosing groups.

The summaries will only include data while subjects are receiving double-blind study drug. Data will not be included for visits after a histologic non-responder at Week 12 had received single-blind (to subject) APT-1011 3 mg BID in Part 2.

Patients who do not have assessments at the respective visits will not be included in the denominator for

percentages.

No statistical testing will be performed.

16.2.3.4. Analysis of Change from baseline in each of Global EoE Symptom Score, Total EREFs Score, Number of Dysphagia Episodes, 7-day EEsAI total score, 7-day EEsAI VDQ and AMS subscores

The change from baseline in each of Global EoE Symptom Score (see section 16.2.1.4), 7-day EEsAI total score (see section 16.2.1.6) and 7-day EEsAI VDQ and AMS subscores (see section 16.2.1.7) will be summarized as continuous variables, and analysis of covariance (ANCOVA) will be used to compare the mean change from baseline in each of the APT-1011 dosing groups to placebo, at the Week 12 visit, with the estimated mean difference, 90% confidence interval of the mean difference and 1-sided p-values provided. This statistical testing at Week 12 is considered descriptive only and no adjustments will be made for multiple comparisons.
The ANCOVA models will include the factors described in section 7.1.

All these change from baseline analyses will only include data while subjects are receiving double-blind study drug. Data will not be not included for visits after a histologic non-responder at Week 12 had received single-blind (to subject) APT-1011 3 mg BID in Part 2.

For the analysis of change from baseline in the number of dysphagia episodes (see section 16.2.1.5) and dysphagia free days , for each dose that meets the criterion for statistical significance on the primary efficacy outcome, a Wilcoxon Rank-Sum test will be used to test whether the reduction in the number of dysphagia episodes from baseline is larger for the APT-1011 dosing group compared to the control.

Holm's step-down procedure (Holm, 1979) will be used to control the overall Type I error at 0.05, where the number of doses considered for inferential testing is equal to the number of doses that meet statistical significance for the primary efficacy analysis.

Holm's step down procedure will be implemented as follows:
Let the number of APT-1011 doses that meet statistical significance in the primary efficacy analysis be represented by d (where d is between 1 and 4 inclusive for this study).

Perform the Wilcoxon-Rank Sum test for each of the d doses vs placebo and rank the resulting p-values in ascending order: $p_1 <= p_i <= p_d$

For each dose the hypothesis to be tested is:

H_0: $Ri(x) \leq R_0(x)$

H1: $Ri(x) > R_0(x)$

where R_i (x) and $R_0(x)$ are the distribution functions for the reduction in number of dysphagia episodes for APT- 1011 dosing group i and placebo respectively.

Step 1: Test the dose with smallest p-value (p_1)

The hypothesis H_0 will be rejected if the one-sided p-value p_1 is $< 0.05/d$

Step 2 : Test the dose with second smallest p-value (p_2)

The hypothesis H_0 will be rejected if the one-sided p-value p_2 is $< 0.05/d-1$

Step 3 : Test the dose with third smallest p-value (p_3)

The hypothesis H_0 will be rejected if the one-sided p-value p_3 is $< 0.05/d-2$

Step 4 : Test the dose with fourth smallest p-value (p_4)

The hypothesis H_0 will be rejected if the one-sided p-value p_4 is $< 0.05/d-3$

Note that depending on the value of d, not all of these steps may be required.

16.2.3.5. Analysis of PGIC

The number and percentage of subjects in each PGIC category (see section 16.2.1.10) will be presented for each of the randomized dosing groups at each post-baseline visit. Data will not be included for visits once a subject has switched to single-blind (to subject) APT-1011 3 mg BID.

No statistical testing will be performed.

16.2.3.6. Analysis of percentage of histologic non-responders at Week 12, Week 26, and Week 52

The frequency and percentage of non-responders will be provided for each week.

Data are not included for visits after histologic non-responders at Week 12 had received single-blind (to subject) APT-1011 3 mg BID in Part 2.

No statistical testing will be performed.

[redacted]

16.4. ADDITIONAL EFFICACY

16.4.1. ADDITIONAL EFFICACY ANALYSES

All analyses of the primary and secondary efficacy variables, including the sensitivity analyses, will be repeated using the ITT population.

17. SAFETY OUTCOMES

All outputs for safety outcomes will be based on the SAF.
Safety will be presented separately for each of Part 1, Part 2 and for Parts 1 and Parts 2 combined.

- For Part 1 all subjects will be included in their double-blind dosing groups.
- For Part 2, histologic responders at Week 12 who continue into Part 2 will be included in their double-blind dosing groups.
- For Part 2, histologic non-responders at Week 12 who continue into Part 2 will be included in a single-blind APT-1011 3.0 mg BID group. Additionally, a summary will be provided for all APT-1011 3.0 mg BID during Part 2, both double-blind and single-blind.

[redacted]

- For Parts 1 and 2 combined, histologic non-responders at Week 12 will be included in their double-blind dosing group (summarizing just their Part 1 safety data) and also in the single-blind APT-1011 3.0 mg BID group (summarizing just their Part 2 safety data). A summary will be provided for all APT-1011 3.0 mg BID during Part 2, both double-blind and single-blind.
- The inclusion of a group combining both double-blind and single-blind APT-1011 3.0 mg should enable a complete summary of the safety of APT-1011 3.0 mg BID in this study. And informal comparisons of original randomized dosing groups will also be possible.

There will be no statistical comparisons between dosing groups for safety data.
Safety data obtained during the placebo run-in period will be listed only.

17.1. ADVERSE EVENTS

Adverse Events (AEs) will be coded using the MedDRA coding dictionary, Version 21.0.

Treatment emergent AEs (TEAEs) are defined as AEs that started or worsened in severity on or after the first dose of study drug. In section 5.0 of the protocol it states: 'the term study drug is used to refer to any blinded medication administered (i.e., any dosage of APT-1011 or placebo)'.

For further defining TEAEs below, study drug will also include single-blind (to subject) APT-1011 3.0 mg BID which histologic non-responders at Week 12 receive in Part 2.

TEAEs will be defined for each of the following periods:

- Part 1 of the study
- Part 2 of the study (with first dose of study drug being the first dose in Part 2)
- Parts 1 and 2 combined

For Parts 1 and 2 combined, histologic non-responders at Week 12 will be summarized for both their double-blind dosing in Part 1 and their single-blind (to subject) dosing in Part 2.

See Appendix 2 for handling of partial dates for AEs. In the case where it is not possible to define an AE as treatment emergent or not, the AE will be classified by the worst case; i.e. treatment emergent.

An overall summary of number of subjects and number of events within each of the categories described in the sub-

section below, will be provided as specified in the templates.

Listings will include TEAEs and Non-TEAEs.

17.1.1. ALL TEAEs

Incidence of TEAEs will be presented by System Organ Class (SOC) and Preferred Term (PT) and also broken down further by maximum severity and relationship to study drug.

17.1.1.1. Severity

Severity is classed as mild/ moderate/ severe (increasing severity). If a subject reports a TEAE more than once within that SOC/ PT, the AE with the worst case severity will be used in the corresponding severity summaries.

17.1.1.2. Relationship to Study Drug

Relationship for AEs, will be indicated by the Investigator as related or not related.

17.1.2. TEAEs LEADING TO STUDY DRUG DISCONTINUATION AND STUDY WITHDRAWAL

TEAEs leading to permanent discontinuation of study drug and sudy withdrawal will be identified as those with an answer of yes to the question '*Did the AE cause the subject to discontinue from the study?*' on the *Adverse Events* page of the eCRF.
For TEAEs leading to discontinuation, summaries of incidence rates (frequencies and percentages) by SOC and PT will be prepared.

17.1.3. SERIOUS ADVERSE EVENTS

Serious AEs (SAEs) are those events recorded as "Serious" on the *Adverse Events* page of the eCRF. A summary of serious TEAEs by SOC and PT will be prepared.

17.1.3.1. Relationship to Study Drug

Relationship for SAEs, will be indicated by the Investigator as related or not related. TESAEs with a missing relationship to study drug will be regarded as related to study drug in summaries.

17.1.4. ADVERSE EVENTS LEADING TO DEATH

TEAEs leading to Death are those events which are recorded as "Fatal" on the *Adverse Events* page of the eCRF. A summary of TEAEs leading to death by SOC and PT will be prepared.

17.1.5. ADVERSE EVENTS OF SPECIAL INTEREST

Adverse events of special interest include oral and esophageal candidiasis and symptoms of adrenal suppression or hypercorticism.
These will be identified from those ticked as yes to the question '*is this an AE of special interest*?' on the *Adverse Events* page of the eCRF.

17.2. LABORATORY EVALUATIONS

Results from the central laboratory will be included in the safety reporting of this study for hematology, serum chemistry and urinalysis. Results will be provided by the central laboratory in SI units. IQVIA Biostatistics will not be performing any conversions to SI units.
A list of hematology, serum chemistry and urinalysis assessments to be included in safety outputs is included in the protocol, sections 6.3.2.1, 6.3.2.2 and 6.3.2.3.

The following summaries will be provided for laboratory data:

- Incidence of clinically significant out of range post-baseline values (for quantitative measurements and categorical measurements)
- Incidence of potential Hy's law (see section 17.2.2) at any time post-baseline.

17.2.1. LABORATORY SPECIFIC DERIVATIONS

Not applicable.

17.2.2. LABORATORY REFERENCE RANGES AND MARKEDLY ABNORMAL CRITERIA

Quantitative laboratory measurements will be compared with the central laboratory reference ranges in SI units and categorized as:

- Low: Below the lower limit of the laboratory reference range.
- Normal: Within the laboratory reference range (upper and lower limit included).
- High: Above the upper limit of the laboratory reference range.

Values outside of the reference ranges and those determined to be clinically significant will be flagged as such in the data listings.

For qualitative and quantitative laboratory measurements, clinically significant values will be identified using the ALERT flag in the central laboratory data provided to IQVIA Biostatistics.

As per section 6.3.2.1 of the protocol, for this study potential Hy's law is defined as alanine transaminase (ALT) or asparate transaminase (AST) ≥3x ULN, and total bilirubin ≥2x ULN.

17.3. ECG EVALUATIONS

Results from the central Electrocardiogram (ECG) review will be included in the reporting of this study.
The following ECG parameters will be listed for this study:

- PR Interval (msec)
- QRS Interval (msec)
- RR Interval (msec)
- QT Interval (msec)
- QTcF Interval (msec)
- QTcB Interval (msec)
- HR (bpm)
- Overall assessment of ECG (Investigator's judgment):
    - Normal
    - Abnormal, Not Clinically Significant (ANCS)
    - Abnormal, Clinically Significant (ACS)

The following summaries will be provided for ECG data:

- Incidence of markedly abnormal criteria at any time post-baseline
- Incidence of clinically significant abnormal ECGs at any time post-baseline

17.3.1. ECG Specific Derivations

All the parameters listed in section 17.3 will be included in data provided by the ECG central reviewer, and will not be derived by IQVIA Biostatistics.

17.3.2. ECG Markedly Abnormal Criteria

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

- Absolute values for QT interval, QTc interval, QTcB interval and QTcF will be classified as:
    - > 450 msec
    - > 480 msec
    - > 500 msec
- Change from Baseline for QT interval, QTc interval, QTcB interval and QTcF will be classified as:
    - >30 msec increase from baseline
    - >60 msec increase from baseline
- Overall assessment of clinical significance (determined by investigator)

These markedly abnormal criteria will be programmed by IQVIA Biostatistics (with the exception of overall clinical significance).

17.4. Vital Signs

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

- Systolic Blood Pressure (SBP) (mmHg)
- Diastolic Blood Pressure (DBP) (mmHg)
- Pulse Rate (bpm)
- Respiratory Rate (breaths/min)
- Temperature (^{0}C)
- Weight (kg)

The following summaries will be provided for vital signs data:

- Baseline and value by visit
- Incidence of markedly abnormal values at any time post-baseline
- Listing of subjects meeting markedly abnormal criteria.

17.4.1. Vital Signs Markedly Abnormal Criteria

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

Variable	Unit	Low	High
SBP	mmHg	≤ 90 mmHg	≥ 180 mmHg
DBP	mmHg	≤ 50 mmHg	≥ 105 mmHg
Heart rate	bpm	≤ 50 bpm	≥ 120 bpm
Respiratory rate	resp/min	≤ 12 resp/min	≥ 25 resp/min
Body temperature	˚C	≤ 35.6 ˚C	≥ 38.3 ˚C
Weight	Kg	≤ 45 kg	>= 120 kg

17.5. Physical Examination

Newly occurring clinically significant physical examination abnormalities will be captured and summarized as AEs. Physical examination data will be listed only.

17.6. OTHER SAFETY ASSESSMENTS

- The number and percentage of subjects with serum cortisol level ≤5 µg/dL (≤138 nmol/L) or abnormal ACTH stimulation test (serum cortisol <16 µg/dL [≤440 nmol/L]) will be presented
- The number and percentage of subjects discontinuing for HPA axis suppression (adrenal suppresion) will be presented
- The number and percntage of subjects discontinued for hypercorticism
- The number and percentage of subjects discontinuing due to abnormal ACTH stimulation test will be presented.

18. REFERENCES

Goria, M.N. 1992. On the forth root transformation of chi-square. Australian Journal of Statistics, 34 (1), 55-64.

Liublinska, V. and Rubin, D.B. Sensitivity analysis for a partially missing binary outcome in a two-arm randomized clinical trial. Statistics in Medicine, 33(24), 4170–4185

Ratitch, B. Lipkovich, I and O'Kelly M, *Combining Analysis Results from Multiply Imputed Categorical Data,* PharmaSUG 2013 - Paper SP03

Rubin, D.B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons.

Wilson, E.B. & Hilferty, M.M. 1931. The distribution of chi-squared. Proceedings of the National Academy of Sciences, Washington, 17, 684–688.

APPENDIX 1. PROGRAMMING CONVENTIONS FOR OUTPUTS

IQVIA OUTPUT CONVENTIONS

The standards below apply for all tables, listings and figures. They provide specifications to guide the statistician or statistical programmer in setting up specifications for programming tables, listings and figures. These standards are used in the absence of customer specific standards.

1. ABBREVIATIONS USED BELOW

ASCII American standard code for information interchange file format
CGM Computer graphics metafile
ODS Output Delivery System
RTF Rich text file format

2. OUTPUT FILE NAMING CONVENTIONS

File names should only consist of uppercase letters, lowercase letters, digits (0 to 9) and underscores. A period should only be used to indicate a separator between the file name and the extension. No spaces, other special characters or punctuation marks are permitted.

As far as possible, output files should be in RTF format, although .DOC files are also permitted.

The program, program log and output file name should reflect the type and number of the statistical output. If this is not possible, then the output name should be at least as descriptive as possible. A prefix can be used to distinguish between a Table, Listing and Figure document ('T' for table, 'L' for listing and 'F' for figure). If there is only 1 digit in the number of the table, listing or figure in the place where 2 digits are possible, a leading zero should be added in the file name to make sorting consistent with the sequence (eg T14_3_01_1.RTF)

4. PAPER SIZE, ORIENTATION AND MARGINS

The size of paper will be Letter for the United States, otherwise A4.
The page orientation should preferably be landscape, but portrait is also permitted.

Margins should provide at least 1 inch (2.54 centimeters) of white space all around the page, regardless of the paper size.
The number of columns per page (linesize) should be 134 for Letter.
The number of rows per page (pagesize) should be 51 for Letter.

5. FONTS

The font type 'Courier New' should be used as a default for tables and listings, with a font size of 8. The font color should be black. No bolding, underlining italics or subscripting should be permitted. Try to avoid using super-scripts, unless absolutely necessary. Single spacing should be used for all text.
Figures should have a default font of "Times Roman", "Helvetica", or "Courier New" .
This can be achieved by using the following options in SAS:

```
goptions
gunit = pct
cback = white
colors = (black)
hby = 2.4
ftext = "TimesRoman"
htext = 2.5
device = cgmof97l
gaccess = gsasfile;
filename gsasfile "….cgm";
```

6. HEADER INFORMATION

Headers should be defined as follows:
• The header should be placed at the top of the page (same place on each page) regardless of the size or orientation of the table or listing
• The customer name and protocol number should appear in row 1, left-aligned
• The output identification number should appear in row 2, centered
• The output title should start in row 3, centered
• The output population should appear in row 4, centered. The population should be spelled out in full, e.g. Intention-to-Treat in preference to ITT.
• Row 5 should be a continuous row of underscores ('_') (the number of underscores should equal the linesize)

• Row 6 should be a blank line
• Mixed case should be used for titles
• The output titles should be designed so that they are arranged consistently through all outputs. For example, content (eg Vital Signs) followed by metric (eg Change from Baseline) e.g. Vital Signs – Change from Baseline.
• Titles should not contain quotation marks or footnote references
• The column headings should be underlined with a row of underscores ('_')
• Column headings spanning more than one column should be underlined and should be centered
• Column headings containing numbers should be centered
• Column headings should be in sentence case
• In general, the population count should appear in the column header in the form "(N=XXX)"
• As a rule, all columns should have column headings.

7. TABLE AND LISTING OUTPUT CONVENTIONS

General:
• The first row in the body of the table or listing should be blank
• The left hand column should start in column 1. No indenting or centering of the output should occur.
• Rounding should be done with the SAS function ROUND.
• Numbers in tables should be rounded, not truncated.
• Alphanumeric output should be left aligned.
• Numbers should be decimal point aligned.
• Whole numbers should be right aligned.
• Text values should be left aligned.
• The first letter of a text entry should be capitalized
• Listings of adverse events, concomitant medications, medical histories etc. should be sorted in chronological order, with earliest adverse event, medication or history coming first.
• The study drug should appear first in tables with treatments as columns
• In general, only present totals (across treatment groups) at baseline/randomization, and do not present them post randomization, unless the customer specifically requests it.
• If possible, include 100% frequencies in the table shell, so that it is clear what the denominator is for percentage calculations.
• All listing outputs should be sorted (preferably by Treatment, Site Number and Subject Number).
• Do not use superscripts and subscripts
• Exponentiation will be expressed using a double asterisk, i.e., mm3 will be written as mm**3.

• All variables that are output in the CRF (which have data present) should appear in the listings, along with all derived data appearing in the corresponding tables
• The width of the entire output should match the linesize

Univariate Statistics:
• Statistics should be presented in the same order across tables (i.e., n, Mean, SD, Median, Minimum, Maximum)
• Table statistics should line up under the N part of the (N=XXX) in the table header. All decimal points should line up. If the minimum and maximum are output on one line as Minimum, Maximum then the comma should line up with the decimal point.
• If the original data has N decimal places, then the summary statistics should have the following decimal places:
Minimum and maximum: N
Mean, median and CV%: N + 1
SD: N + 2

Frequencies and percentages (n and %):
• Percent values should be reported inside parentheses, with one space between the count and the left parenthesis of the percentage. Parentheses should be justified to accept a maximum of 100.0 as a value and padded with blank space if the percent is less than 100.0. An example is given below:
77 (100.0%)
50 (64.9%)
0 (0.0%)
• Percentages will be reported to one decimal place, except percents <100.0% but >99.9% will be presented as '>99.9%' (e.g., 99.99% is presented as >99.9%); and percents < 0.1% will be presented as '<0.1%' (e.g., 0.08% is presented as <0.1%). Rounding will be applied after the <0.1% and >99.9% rule.
Eg (<0.1%)
(6.8%)
(>99.9%)
Percentages may be reported to 0 decimal places as appropriate (for example, where the denominator is relatively small).
• Where counts are zero, percentages of 0.0% should appear in the output.

Confidence Intervals:
• As a rule confidence intervals are output to one place more than the raw data, and standard deviations and standard errors to two places more than the raw data

• Confidence intervals should be justified so that parentheses displayed on consecutive lines of a table “line up”.
• Boundary values of confidence intervals should be separated by a comma.
• Boundary values should be padded as necessary to accept negative values and to allow alignment of the decimal place.
• An example is given below:
(-0.12, -0.10)
(9.54, 12.91)

P-values:
• P-values should be reported to three decimal places, except values <1.000 but >0.999 will be presented as ‘>0.999’ (e.g., 0.9998 is presented as >0.999); and values <0.001 will be presented as ‘<0.001’ (e.g., 0.0009 is presented as <0.001). Rounding will be applied after the <0.001 and >0.999 rule

Ratios:
• Ratios should be reported to one more decimal place than the original data.

Spacing:
• There must be a minimum of 1 blank space between columns (preferably 2)

Denominators:
• If a different count other than the population count is used for a denominator (within the table) to calculate percentages, there should be a row in the table that identifies that number “n”.
• Alternatively, a footnote should be included in each table with percentages to indicate the denominator for percentages.

Missing values
• A “0” should be used to indicate a zero frequency.
• A blank will be used to indicate missing data in an end-of-text table or subject listing.

8. FIGURE OUTPUT CONVENTIONS

• Figures should be provided in RTF files using the SAS Output Delivery System (ODS), as Computer Graphics Metafile (CGM) formatted graphical output generated by SAS.
• The CGM file itself should contain the title or footer.

• The image should be clear and of high quality when viewed in the Word document, and when printed.
• In general, boxes around the figures should be used.

9. FOOTNOTE INFORMATION

Footers should be defined as follows:
• A continuous line of underscores ('_') will follow the body of the table or listing prior to any footnotes at the bottom of the page
• Table footnotes should be defined using compute statements in the proc report, and should appear directly after the body of the table
• The program path and name and version number (if applicable) should appear as the last footnote at the bottom of the page
• The date/time stamp should also appear in the last footnote at the bottom of the page right aligned
• Footnotes should be left-aligned.
• Footnotes should be in sentence case.
• Only "typewriter" symbols are permitted – eg "*", "$", "#", "@", "&" and "+".
• The choice of footnote symbols should be consistent. E.g. if you have the footnote "# indicates last observation carried forward" for one table, the same symbol and footnote should indicate LOCF for all tables.
• The page identification in the format Page X of Y (where Y is the total number of pages for the output) should appear in the first header, right aligned

Ordering of footnotes should be as follows:

1.) Source data listing reference, if necessary
2.) Abbreviations and definitions
3.) Formulae
4.) P-value significance footnote
5.) Symbols
6.) Specific notes
• Common notes from table to table should appear in the same order.
• The symbols should appear in the same order as what they are defined in the table or listing, from left to right.

10. PROGRAMMING INSTRUCTIONS

Programming instructions must appear in blue font at the end of each table or listing shell. Programming instructions, where necessary, should follow the table or listing shells in blue font, beginning with the words "Programming Note" followed by a colon. These include notes on the output, reminders of how to handle missing values, repeat shells for similar tables etc.

DATES & TIMES

Depending on data available, dates and times will take the form yyyy-mm-ddThh:mm:ss.

SPELLING FORMAT

English US will be used.

PRESENTATION OF DOSING GROUPS

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

Treatment Group	**For Tables and Graphs**	**For Listings (include if different to tables)**
APT-1011 3.0 mg BID	APT-1011 3 mg BID	3 mg BID
APT-1011 3.0 mg HS	APT-1011 3 mg HS	3 mg HS
APT-1011 1.5 mg BID	APT-1011 1.5 mg BID	1.5 mg BID
APT-1011 1.5 mg HS	APT-1011 1.5 mg HS	1.5 mg HS
Placebo	Placebo	Placebo
Non-Responders at Week 12	Single-Blind APT-1011 3 mg BID	Randomized Treatment Group with * concatenated for relevant data
APT-1011 3.0 mg BID and Non-Responders at Week 12 (for Safety Outputs)	All APT-1011 3 mg BID (with footnotes explaining this)	Not for use in listings.

Treatment Group	For Tables and Graphs	For Listings (include if different to tables)
Not Randomized	N/A	Not Randomized

PRESENTATION OF VISITS

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

Long Name (default)	Short Name For Use in Outputs
Visit 1 (Screening)	Screening
Derived Baseline per Parameter	Baseline (for tables only)
Visit 2 (4-week Baseline Symptom Assessment)	Visit 2
Visit 3	Rand.
Visit 4 (Week 4)	Week 4
Visit 5 (Week 8)	Week 8
Visit 6 (Week 12)	Week 12
Visit 7 (Week 14)	Week 14
Visit 8 (Week 18)	Week 18
Visit 9 (Week 22)	Week 22
Visit 10 (Week 26)	Week 26
Visit 11 (Week 28)	Week 28
Visit 12 (Week 36)	Week 36
Visit 13 (Week 44)	Week 44

Long Name (default)	Short Name For Use in Outputs
Visit 14 (Week 52)	Week 52
Unscheduled	Unsch
End of Treatment	EOT

LISTINGS

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

- randomized dosing group (or treatment received if it's a safety output), first by active dose [by descending dose group] and then placebo
- center-subject ID,
- date (where applicable),
- For listings where non-randomized subjects are included, these will appear in a category after the randomized dosing groups labeled 'Not Randomized'.
- For visits or safety data where the data were captured for non-responders at Week 12 after they had received single-blind APT-1011 3mg BID in Part 2, the visit or the event will be annotated with a * (e.g. Week 28*) for that record with a footnote explaining that * denotes information after the subject received single-blind (to the subject) APT-1011 3 mg BID in Part 2.

APPENDIX 2. PARTIAL DATE CONVENTIONS

Imputed dates will NOT be presented in the listings.

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS:

START DATE	STOP DATE	ACTION
Known	Known, Partial or Missing	If start date < study drug start date, then not TEAE If start date >= study drug start date, then TEAE

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

ALGORITHM FOR ASSIGNMENT OF TREATMENT EMERGENT ADVERSE EVENTS TO STUDY PARTS:

The rules below will be applied to AEs which were defined as TEAEs using the table above.

START DATE	STOP DATE	ACTION
Known	Known, Partial or Missing	If start date >= Part 2 study drug start date assign to Part 2 Otherwise: assign to Part 1
Partial	Known	If stop date < Part 2 study drug start date assign to Part 1 Otherwise: If known components of start date are such that the start date is definitely >= Part 2 study drug start date, assign to Part 2 or if known components of start date are such that the start date is definitely < Part 2 study drug start date, assign to Part 1. Or if known components of start date are such that the start date might be >= Part 2 study drug start date, assign to Part 1 and Part 2
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If imputed stop date < Part 2 study drug start date assign to Part 1. Otherwise: If known components of start date are such that the start date is definitely >= Part 2 study drug start date, assign to Part 2 Or if known components of start date are such that the start date is definitely < Part 2 study drug start date, assign to Part 1. Or if known components of start date are such that the start date might be >= Part 2 study drug start date, assign to Part 1 and Part 2
	Missing	If known components of start date are such that the start date is definitely >= Part 2 study drug start date, assign to Part 2

START DATE	STOP DATE	ACTION
		Or if known components of start date are such that the start date is definitely < Part 2 study drug start date, assign to Part 1. Or if known components of start date are such that the start date might be >= Part 2 study drug start date, assign to Part 1 and Part 2
Missing	Known	If stop date < Part 2 study drug start date assign to Part 1 and Part 2
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If imputed stop date < Part 2 study drug start date assign to Part 1 Otherwise: assign to Part 1 and Part 2
	Missing	Assign to Part 1 and Part 2

ALGORITHM FOR PRIOR / CONCOMITANT MEDICATIONS:

START DATE	STOP DATE	ACTION
Known	Known	If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment, assign as concomitant If stop date >= study drug start date and start date > end of treatment, assign as post study

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

START DATE	STOP DATE	ACTION
	Missing	Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown), then: If stop date is missing could never be assumed a prior medication If start date <= end of treatment, assign as concomitant If start date > end of treatment, assign as post treatment
Missing	Known	If stop date < study drug start date, assign as prior If stop date >= study drugstart date, assign as concomitant Cannot be assigned as 'post treatment'
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date, assign as concomitant Cannot be assigned as 'post treatment'
	Missing	Assign as concomitant