# STATISTICAL ANALYSIS PLAN

ARC007
Real-World AR101 Market-Supporting Experience Study in Peanut-Allergic Children Ages 4 to 17 Years (RAMSES)

Version 2.0 – 27 Sep 2018
Reference Numbers: NCT03126227

Aimmune Therapeutics, Inc.
8000 Marina Blvd, Suite 300
Brisbane, CA 94005
United States

| | |
|---|---|
| **Protocol Title:** | Real-World AR101 Market-Supporting Experience Study in Peanut-Allergic Children Ages 4 to 17 Years (RAMSES) |
| **Protocol Identifier:** | ARC007 |
| **Investigational Product:** | AR101; Peanut Allergen (*Arachis hypogaea*) (Formerly Characterized Peanut Allergen or CPNA) |
| **Protocol Version and Date:** | Amendment 2.0, 27 Sep 2017 |
| **Sponsor:** | Aimmune Therapeutics, Inc.<br>8000 Marina Blvd, Suite 300<br>Brisbane, CA 94005<br>United States |
| **Authors:** | Trinh Pham, MS<br>Laura Holland, MS |
| **SAP Version and Date:** | Final Version 2.0; 27 September 2018 |

## STATISTICAL ANALYSIS PLAN APPROVAL

## PROTOCOL ARC007

28 SEP 2018

**Kari Brown, MD**
**Medical Director, Clinical Science**
**Aimmune Therapeutics, Inc.**

**Date**

28 SEP 2018

**Rezi Zawadzki, DrPH**
**Vice President, Biometrics**
**Aimmune Therapeutics, Inc.**

**Date**

28 SEP 2018

**Laura Holland, MS**
**Biostatistician I**
**Precision for Medicine, Oncology and Rare Disease**

**Date**

28 SEP 2018

**Brian Beus, MS**
**Principal Biostatistician**
**Precision for Medicine, Oncology and Rare Disease**

**Date**

# TABLE OF CONTENTS

## LIST OF ABBREVIATIONS AND TERMS

| Abbreviation | Description |
|---|---|
| ACT | Asthma Control Test |
| ATC | Anatomical therapeutic chemical |
| CODIT | Characterized oral desensitization immunotherapy |
| CoFAR | Consortium of Food Allergy Research |
| CRF | Case report form |
| CTCAE | Common Terminology Criteria for Adverse Events |
| DSMC | Data safety monitoring committee |
| FAIM | Food allergy independent measure |
| FAQLQ | Food allergy-related quality of life questionnaire |
| GI | Gastrointestinal |
| IgE | Immunoglobulin E |
| IgG4 | Immunoglobulin G subclass 4 |
| MedDRA | Medical Dictionary for Regulatory Activities |
| OIT | Oral immunotherapy |
| PERF | Peak expiratory flow rate |
| PRN | Pro re nata (as needed) |
| PS | Peanut-specific |
| QD | Once a day |
| SAP | Statistical analysis plan |
| SOC | System organ class |
| SPT | Skin prick test |
| SI | Subject identification |
| TEAE | Treatment-emergent adverse event |
| TNSS | Total Nasal Symptom Score |
| WHO-DDE | World Health Organization Drug Dictionary Enhanced |

# 1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to ensure that the statistical methodologies employed and the final analysis summary tables, figures, and data listings produced will be complete and appropriate to allow valid conclusions regarding the study objectives.

Aimmune Therapeutics, Inc. prepared the SAP, and Precision for Medicine, Oncology and Rare Disease ("Precision") will perform the analyses and be responsible for the production and quality control of the analysis summary tables, figures, and listings.

This SAP describes the analysis of data collected from ARC007 only. Subjects who received AR101 in study ARC007 and completed the study will have the option to continue treatment with AR101 300 mg once daily (QD) for the duration of the 6-month maintenance period in study ARC011; therefore, a separate SAP or addendum to this SAP will be written to provide a complete safety profile for subjects from initial escalation to the end of the maintenance period for those subjects.

# 2 STUDY OVERVIEW

ARC007 is a multicenter, randomized, double-blind, placebo-controlled safety study of AR101 using the characterized oral desensitization immunotherapy (CODIT™) regimen in peanut-allergic children aged 4 to 17 years, inclusive. The study consists of screening and a double-blind treatment period that includes an initial 2-day escalation and up-dosing (approximately 20-40 weeks). After completion of up-dosing, all study exit procedures, and treatment unblinding, subjects who received AR101 may participate in study ARC011 to receive open-label maintenance treatment with AR101. Subjects who received placebo will be offered up-dosing and maintenance treatment with AR101 in study ARC008.

For additional details on study design and procedures, refer to protocol Section 3 and Section 6.

# 3 STUDY OBJECTIVES

## 3.1 Primary Objective

- The primary objective is to assess the safety and tolerability of AR101 in a CODIT regimen for approximately 6 months of double-blind treatment in peanut-allergic children aged 4 to 17 years, inclusive.

### 3.2 Secondary Objectives

- To characterize the frequency of all treatment-related adverse events by study period
- To assess the effect of AR101 on asthma control and immune parameters

## 4 STUDY ENDPOINTS

Study ARC007 is a safety study; there are no efficacy analyses, and all study endpoints are considered safety endpoints.

### 4.1 Primary Endpoint

The primary endpoint is the frequency of treatment-emergent adverse events, including serious adverse events, during the overall study period.

### 4.2 Secondary Endpoints

Secondary endpoints are as follows:

- Frequency of premature discontinuation of dosing due to adverse events
- Frequency of premature discontinuation of dosing due to chronic/recurrent gastrointestinal (GI) adverse events
- In subjects who discontinue study early due to chronic/recurrent GI AEs, proportion whose ongoing GI adverse events resolved prior to 2, between 2 and 4, between 4 and 12, and ≥ 12 weeks following cessation of dosing
- Frequency of allergic reaction (hypersensitivity) adverse events occurring during up-dosing, normalized for duration of treatment
- Frequency of anaphylaxis as defined in the protocol
- Frequency of use of epinephrine as a rescue medication
- Frequency of ingestion of peanut (not study product) and other allergenic foods and severity of any resultant reactions
- Assessment of asthma control using the Asthma Control Test (ACT) questionnaire and frequency of use of asthma rescue medication (short acting beta-agonists) in subjects with asthma
- Frequency of adverse events that lead to early withdrawal

### 4.3 Exploratory Endpoints

The exploratory endpoints are as follows:

- Change in peanut-specific and peanut component-specific serum immunoglobulin (Ig) E and IgG4 levels
- Changes in peanut skin prick test (SPT) mean wheal diameter
- Changes in Total Nasal Symptom Score (TNSS) in subjects with allergic rhinitis

- Changes in scores of the Food Allergy Quality of Life Questionnaire (FAQLQ) and Food Allergy Independent Measure (FAIM)

# 5 SAMPLE SIZE CONSIDERATIONS

ARC007 is a safety study designed to add critical placebo-controlled data to the ongoing safety database of AR101. As such, there are no efficacy endpoints in the study, no specific efficacy-related hypotheses to be tested, and no prospective sample size calculations performed related to the power to detect a prespecified treatment effect size. A sample size of up to 500 subjects randomized in a 2:1 ratio to AR101 or placebo, along with subjects enrolled in other studies in the clinical program was to provide a sufficient number of subjects to fulfill the regulatory requirement for data on at least 600 subjects treated with AR101 at 300 mg/day for 6 months. Randomization was stratified by age group (4-11 years and 12-17 years).

# 6 ANALYSIS POPULATION

## 6.1 Safety Population

The safety population is defined as all subjects who received at least 1 dose of randomized study treatment in ARC007. Subjects will be analyzed according to treatment received.

The safety population of pediatric subjects aged 4 to 17 years will be the population for all analyses unless otherwise specified.

# 7 DEFINITIONS, DERIVATIONS, AND CONVENTIONS

## 7.1 Definitions and Derivations

- Study day is calculated as (assessment date – first dose date + 1) for assessments and visits performed on or after the first dose date, and (assessment date – first dose date) for assessments and visits prior to the first dose date.

- Baseline is defined as the last nonmissing value prior to the first dose of randomized study treatment in ARC007.

- Change from baseline is calculated as the observed value after the first dose – baseline value.

- The screening period is defined as the time beginning with the date and time of informed consent through the day prior to randomization.

- Initial escalation is defined as the time beginning with the date and time of the first dose of randomized study product at the study site and ending with the date of the last dose of randomized study product taken prior to up-dosing.

- Up-dosing is defined as the time beginning with the date and time of the first dose of study product at 3 mg at home, and ending with the date and time of first dose at

300 mg at the study site. Up-dosing will be approximately 20 weeks in duration, but may be extended to a maximum of 40 weeks to accommodate dose reductions and re-escalations, if necessary.

- A 300 mg QD period will be defined as the time beginning with the date and time of the first dose of study product at 300 mg at home and ending with the date of the last dose of study product. This period will be ideally approximately 2 weeks in duration at the end of ARC007, but dosing may continue up to a maximum of 48 weeks after day 1 to ensure tolerability of 300 mg prior to exiting ARC007. Some subjects continued to receive 300 mg doses in ARC007 after the expected rollover date while waiting for activation of ARC008 or ARC011 at the study site. The 300 mg doses as a result of the rollover delay will be included in the 300 mg QD period.

## 7.2 Programming Conventions

Unless stated otherwise, the term "descriptive statistics" refers to the number of subjects (n), mean, median, standard deviation, minimum, and maximum for continuous variables and frequencies and percentages for categorical variables.

Unless specified otherwise, the denominator for percentages for categorical data will be based on the number of subjects or observations with nonmissing data appropriate for summary purposes. The denominator for percentages for incidence data (such as adverse events) will be based on the number of subjects in the analysis population "at risk".

Minimum and maximum values will be presented at the precision of the original value, means, medians will be rounded to 1 decimal place greater than the precision of the original value, standard deviations and standard errors will be rounded to 2 decimal places greater than the precision of the original value. Percentages will be rounded to 1 decimal place. Percentages that round down to 0 or up to 100% will be displayed as "< 0.1%" and "> 99.9%", respectively.

All summary tables will be presented by treatment group displayed as AR101 and placebo. For disposition, demographic, and other summaries of baseline and history data, a total column for both treatment groups combined will be included.

All relevant data collected in the database will be included in data listings and sorted by treatment group, subject number, and visit and time point as appropriate. The treatment group will be displayed by AR101 then placebo.

Unscheduled visits will be listed but not included in by-visit summaries. Results from unscheduled visits may be used as baseline values and for other derivations not tied to visit names (for example, unscheduled visits are included in the determination of worst postbaseline values for laboratory shift tables).

### 7.2.1 Rules for Missing Data

All adverse events with partial/missing dates and times will be considered treatment-emergent adverse events unless a partial date clearly indicates that it occurred

before the first dose of study treatment. All therapies with partial or missing dates and times recorded on the concomitant medication or non-drug therapy case report form pages will be considered concomitant unless a partial stop date and time clearly indicates it was stopped prior to the first dose of study treatment. Start and stop dates will be imputed when partial dates are present as needed to determine treatment-emergent events and concomitant medications. No imputation will be done for a completely missing start/stop date or for subjects who did not receive study treatment.

Start dates with a missing day but with month and year populated will be imputed such that:

- If the provided month and year match the month and year for that subject's first dose date, then the first dose date will be used
- In all other cases the first day of the month will be used with the provided month and year

Start dates with a missing day and month with year populated will be imputed such that:

- If the provided year matches the year for that subject's first dose date, then the first dose date will be used
- In all other cases, January 1 will be used with the provided year

Stop dates will be imputed as follows:

- Missing day with a provided year and month will use the last day of the month
- Missing day and month with provided year will use December 31

If the imputed stop date is greater than the last study date for the subject, then the imputed date will be replaced with the last known subject date.

The reported date of the reaction on the peanut allergy history case report form page and date of diagnosis of peanut allergy will be imputed when the month or day is missing as follows:

- Missing day is set to 1 if the same year and month as the informed consent date. Otherwise it is set to 15
- Missing month and day are set to January 1 if the same year as the informed consent date. Otherwise it is set to July 1.

No imputations will be made for other missing data unless specified otherwise.

### 7.2.2 Visit Windows

All information will be listed, summarized, and analyzed according to the nominal visit time point, study period, or dose. No visit windowing will be performed.

## 8 TIMING OF ANALYSIS

Analyses for data safety and monitoring committee (DSMC) review were planned to occur approximately every 3 months to monitor safety.

The final analysis of safety will be performed after all subjects complete exit/early termination visit assessments in ARC007. Subjects who received AR101 and completed

ARC007 have the option to participate in the ARC011, an open-label safety extension study to evaluate the safety and tolerability of AR101 300 mg/day maintenance treatment for up to 6 months.

# 9 STATISTICAL METHODS

## 9.1 Subject Disposition

The number of subjects randomized and the number and percentage of subjects in the safety population, completed and discontinued, and entered each study period will be summarized for all randomized subjects by age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo, and total).

Reasons for discontinuation from the study will be summarized. In addition, subjects who require follow-up due to chronic/recurrent GI adverse events will be summarized.

Subject completion status, date of study completion/discontinuation, study treatment discontinuation, and reason for discontinuation will be listed based on the information collected on the case report form.

Inclusion and exclusion eligibility will be listed separately.

## 9.2 Protocol Deviations

A protocol violation (ie, major protocol deviation) is a deviation from the protocol approved by the institutional review board or ethics committee that may affect the subject's rights, safety, or well-being and/or the completeness, accuracy, and reliability of the study data. In addition, protocol violations include willful or knowing breaches of human subject protection regulations or policies; any action that is inconsistent with medical and ethical principles; and a serious or continuing noncompliance with federal, state, local, or institutional human subject protection regulations, policies, or procedures.

A nonmajor protocol deviation is any change, divergence, or departure from the study design or procedures that does not have a major impact on the subject's rights, safety, or well-being or the completeness, accuracy, and reliability of the study data.

All protocol deviations will be reported in the data system on a specific case report form according to the following categories:

- Inclusion criteria
- Exclusion criteria
- Received incorrect study treatment
- Randomization issue/ randomized to wrong stratum
- Informed consent form
- Serious adverse event not reported
- Visit out of window
- Missed study visit
- Procedure not per protocol

- Prohibited concomitant medication
- Laboratory sample missed
- Study Product compliance
- Other (with free text field to record details)

Protocol deviations will be reviewed and their categorization as major or minor will be determined before database lock. Major protocol deviations will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo, and total). All protocol deviations, both major and nonmajor, will be listed for all randomized subjects.

## 9.3 Demographics and Baseline Characteristics

Demographic and baseline characteristics will be listed for all randomized subjects.

Summary statistics for demographics and baseline characteristics will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo, and total). Demographic data will include age, race, ethnicity, sex, body weight, height and body mass index (BMI). Baseline characteristics include total IgE, peanut-specific (ps)-IgE, ps-IgG4, ps-IgE/IgG4 ratio, results from SPT, childbearing potential, and history of asthma.

Age will be calculated relative to date of informed consent for ARC007, as follows:

- If the month and day portion of the informed consent date is prior to the month and day portion of the birthdate, age will be calculated as the year of informed consent minus the year of birth, minus 1;

- If the month and day portion of the informed consent date is on or after the month and day portion of the birthdate, age will be calculated as the year of informed consent minus the year of birth.

## 9.4 Disease Characteristics and Previous Therapies

### 9.4.1 Peanut Allergy History

Peanut allergy history will be listed for all randomized subjects.

The duration of peanut allergy (months since peanut allergy diagnosis), months since qualifying allergic reaction to peanut, medication given for qualifying allergic reaction to peanut, number of anaphylactic reactions to peanut in lifetime, months since most recent reaction to peanut, and the allergy symptoms experienced during the most recent peanut exposure will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo, and total).

The reported date of the reaction and date of diagnosis of peanut allergy will be imputed based on the logic in Section 7.2.1.

### 9.4.2 Nonpeanut Allergy History

All nonpeanut allergy history will be listed for all randomized subjects. The presence of nonpeanut allergy history and type of allergy will also be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo, and total).

### 9.4.3 Other Medical History

Medical history will be listed by subject and body system for all randomized subjects. Subjects with abnormal medical history events will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo, and total) and by the Medical Dictionary for Regulatory Activities (MedDRA) system organ class (SOC) version 19.1 and preferred term.

## 9.5 Extent of Exposure and Compliance

Study treatment exposure will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo) and study period (initial escalation, up-dosing, 300 mg QD, and overall).

First and last dose dates for each study period will be identified as follows:

| Study Period | First Dose Date | Last Dose Date |
|---|---|---|
| Initial escalation | Date of first dose at the study site on initial escalation day 1 | Date of last dose at initial escalation day 1 or day 2 at the study site |
| Up-dosing | The day following the date study product was dispensed at initial escalation day 2 | Date of first dose of 300 mg at the end of the up-dosing 300 mg visit if dose is tolerated and subject continues to at-home 300 mg QD dosing. If first in-clinic 300 mg dose is not tolerated and dose is immediately reduced, the last dose date of up-dosing will be when the first in-clinic 300 mg dose is tolerated and subject continues to at-home 300 mg QD dosing. For subjects who do not reach the 300 mg dose, the latest of the last dose during up-dosing at the study site and the last dose at home |
| 300 mg QD | Date of first dose of study product of 300 mg at home. This is defined as the day following the first in-clinic visit where the 300 mg dose was tolerated and 300 mg study product was first dispensed. | Date of last dose of study product for subjects who enter 300 mg QD study period |
| Overall | Date of first dose at the study site on initial escalation day 1 | Last dose of study product prior to exit visit in ARC007 |

The total amount of study product consumed will be calculated as the sum of doses at the study site plus the sum of doses taken at home as recorded in the dosing diary.

The following calculations of study product exposure will be made and summarized:

- Duration of exposure (in days and in months): calculated as the date of the last dose of study product minus the date of the first dose of study product plus 1 during the study period, except for initial escalation, the duration of exposure will only be 1 or 2 days depending on whether study product was taken on initial escalation day 1 and initial escalation day 2. Duration of exposure will be summarized using descriptive statistics for continuous endpoints and categorically by 28-day increments for the overall treatment period: ≤ 28 days, 29–56 days, … 337–364 days, and > 364 days.

- Total dose consumed (mg): calculated as the cumulative sum of all doses taken during the study period.

- Average dose per day (mg): calculated as the total dose consumed divided by the number of days during the study period.

- Number of unsuccessful dose increases: where an unsuccessful dose increase is defined as a single dose at the study site at a higher dose level followed by an immediate return to the previous dose level or a lower dose level.

- Number of dose reductions: where a dose reduction is defined as any decrease in dose level that does not qualify as an unsuccessful dose increase.

- Maximum dose achieved (mg/day): summarized using descriptive statistics for continuous endpoints and categorically using all possible dose levels: 0.5, 1, 1.5, 3, 6, 12, 20, 40, 80, 120, 160, 200, 240, or 300 mg/day.

- Time to 300 mg dosing and time to 80 mg dosing for the overall treatment period using Kaplan-Meier methodology. Time will be calculated as date of the first 300 mg (or 80 mg) dose minus the first dose date +1. Subjects who do not reach the specified dose will be censored at the date of their last study product dose.

The nonmissing valid diary entries will be used to estimate dosing compliance at home. The following measures of compliance with dosing at home will be calculated:

- Total number of planned dosing days at home: calculated as the number of days where a valid diary entry was made, but excluding entries where a dose was missed because of investigator orders.

- Percentage of planned dosing days where a full or partial dose was consumed.

- Percentage of planned dosing days where a full dose was consumed.

- Percentage of planned dosing days where a partial dose was consumed.

- Percentage of planned dosing days where a dose was missed.

Data for dosing at home will be listed. Daily diary records, including date and time, whether a full or partial dose was consumed (or the dose was missed), and the reason for partial or missed doses will be listed. If any dose-related allergy symptoms are present, the symptoms and medications to treat the symptoms will also be listed.

### 9.6 Prior, Concomitant, and Rescue Medications and Therapies

All medications recorded on the concomitant medications case report form page will be coded using the World Health Organization Drug Dictionary Enhanced (WHO-DDE), September 2016 version. Medications will be listed and summarized by Anatomical Therapeutic Chemical (ATC) level 4 and generic name. All prior and concomitant medications, all rescue medications, and epinephrine medications will be listed separately.

Prior medications are defined as those taken prior to the date of the first dose of study product on day 1 (ie, medication end date is prior to the date of first dose of study product).

Concomitant medications are medications taken any time during the active treatment period. Concomitant medications also include medications recorded for which dosing began after the last dose of randomized study treatment. As needed (PRN) medications taken during the active treatment period will be considered concomitant medications. If it cannot be determined whether a medication was received prior to the start of study product dosing due to partial or missing medication start and/or end dates, it will be considered a concomitant medication.

Rescue medications are any medication used to treat symptoms of an acute allergic reaction as identified on the concomitant medication case report form. Each use of a rescue medication during study treatment should be associated with a corresponding adverse event. Prior medications and concomitant medications including rescue medications will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), ATC class, and generic name. Subjects will be counted no more than once per generic name and no more than once per ATC level 4 in the summary.

Rescue medications will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), ATC class and generic name for the following study periods: initial escalation, up-dosing, 300 mg QD, and overall.

Concomitant non-drug therapies will be listed by subject.

### 9.7 Efficacy Analyses

ARC007 is a study to assess safety and tolerability of AR101. As such, all study endpoints are considered safety endpoints and no efficacy analyses will be performed.

### 9.8 Safety Analyses

Safety will be assessed by all the endpoints defined in Section 4.

Safety data will be summarized descriptively and the safety population will be used for all summaries of safety parameters, unless otherwise noted. Safety listings will include all randomized subjects sorted by treatment group (AR101 then placebo).

### 9.8.1 Adverse Events

All adverse events recorded on the adverse event case report form, symptoms recorded on the allergic adverse events case report form, and symptoms recorded on the in-clinic dosing case report will be summarized together.

If symptoms are recorded as part of an anaphylaxis reaction as reported on the Allergic AE or in-clinic dosing form, only the single anaphylaxis reaction event will be summarized and not the individual symptoms.

All reported adverse events will be classified into SOC and preferred term using MedDRA version 19.1.

Treatment-emergent adverse events are defined as those adverse events with onset after the first dose of study product. Adverse events with onset prior to first dose of study product will be included in subject listings, but not summarized. Treatment-emergent adverse events and non-treatment-emergent adverse events will be listed separately. Treatment-emergent adverse events will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), and by study period as follows:

- Initial escalation: All events beginning after the first dose of study product on day 1 and prior to the first administration of study product at home. To avoid gaps between study periods, the first administration of study product at home is assumed to be the day after at-home study product was dispensed at Initial Escalation Day 2.

- Up-dosing: All events beginning after the first administration of study product at home during up-dosing and prior to the first dose of 300 mg at home. To avoid gaps between study periods, the first administration of 300 mg at-home dose is assumed to be the day after the clinic visit where 300 mg dose was first tolerated and at-home 300 mg study product was dispensed.

- 300 mg QD: All events beginning after the first dose of 300 mg at home and prior to exit from the study.

- Overall: Across initial escalation, up-dosing, and 300 mg QD periods (first dose to last dose).

Summaries displayed by SOC and preferred terms will be ordered by descending incidence of SOC and preferred term within each SOC. Summaries displayed by preferred term only will be ordered by descending incidence of preferred term. Summaries of the following types will be presented:

- Overall summary of number of unique treatment-emergent adverse events and treatment-emergent serious adverse events, subject incidence of treatment-emergent adverse events and treatment-emergent serious adverse events meeting various criteria, and exposure-adjusted incidence rates of treatment-emergent adverse events and treatment-emergent serious adverse events meeting various criteria, where exposure incidence rates are defined as the total number of events divided by the total number of subject-years at risk; the total number of subject-years at risk is defined as total number of days for all subjects in the study period divided by 365.25.

- Subject incidence of treatment-emergent adverse events by SOC and preferred term.

- Subject incidence of treatment-emergent adverse events by preferred term.

- Exposure-adjusted event rates for the most frequent treatment-emergent adverse events (ie, treatment-emergent adverse events occurring in ≥ 5% of the safety population) by preferred term.

- Subject incidence of treatment-emergent adverse events by severity grade, SOC, and preferred term.

- Subject incidence of treatment-emergent adverse events related to study product by SOC and preferred term.

- Exposure-adjusted event rates for the most frequent treatment-emergent adverse events related to study product (ie, related treatment-emergent adverse events occurring in ≥ 5% of the safety population) by preferred term.

- Subject incidence of grade ≥ 3 severity treatment-emergent adverse events by SOC and preferred term.

- Subject incidence of grade ≥ 3 severity treatment-emergent adverse events related to study product by SOC and preferred term.

- Subject incidence of serious adverse events by SOC and preferred term.

- Subject incidence of treatment-emergent adverse events leading to discontinuation of study product by SOC and preferred term.

- Subject incidence of treatment-emergent adverse events associated with the use of epinephrine by SOC and preferred term.

- Subject incidence of hypersensitivity treatment-emergent adverse events by SOC and preferred term, where hypersensitivity treatment-emergent adverse events are those treatment-emergent adverse events from the allergic adverse event and in-clinic dosing case report forms.

- Subject incidence of hypersensitivity treatment-emergent adverse events related to study product by SOC and preferred term.

- Exposure-adjusted event rates for hypersensitivity treatment-emergent adverse events by preferred term.

- Subject incidence of treatment-emergent adverse events associated with a non-study product food allergen exposure by SOC, and preferred term.

- Subject incidence of treatment-emergent adverse events leading to early discontinuation by SOC and preferred term.

- Subject incidence of treatment-emergent adverse events with onset < 90 minutes after study product dosing at the study site by SOC and preferred term.

At each level of summarization (eg, any adverse event, SOC, and preferred term), subjects experiencing more than 1 treatment-emergent adverse event will be counted only once within each study period. In the summary of treatment-emergent adverse events by severity grade, subjects will be counted once at the highest severity reported at each level of summarization. Adverse event data will be presented in data listings by age group, treatment group, subject, study period, and event. Treatment-emergent adverse events; serious adverse events; severe, life-threatening, or fatal adverse events; and adverse events leading to discontinuation, reduction, or interruption of the study product will be presented in separate data listings.

### 9.8.2 GI Adverse Events

Of subjects who discontinue due to chronic/recurrent GI adverse events, GI adverse events will be summarized by preferred term, study period, age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo). Additionally, any of these subjects whose GI adverse events were ongoing will be further summarized as having events resolved prior to 2, between 2 and 4, between 4 and 12, and ≥ 12 weeks following cessation of dosing. Summaries will be provided using the safety population by age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo).

All reported GI treatment-emergent adverse events for subjects who discontinue the study early due to chronic/recurrent GI AEs will be listed by age group, treatment group, subject, study period, and event and will include weeks to resolution following cessation of dosing for events that were ongoing.

### 9.8.3 Anaphylaxis

All reported anaphylactic reactions will be listed by age group, treatment group, and subject.

Each anaphylactic reaction will be identified by the following triggers:

- Study product

- Peanut or peanut-containing food
- Other food allergen
- Medication
- Insect sting
- Environmental allergen(s)
- Other

Anaphylactic reactions will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), and treatment group (AR101, placebo) for the following study periods:

- Overall (including initial escalation, up-dosing, 300 mg QD)
- Initial escalation
- Up-dosing
- 300 mg QD

The number of anaphylactic reactions, the number of anaphylactic reactions by trigger (study product, peanut or peanut-containing food, other food allergen, medication, insect sting, environmental allergen, and other), the number and percent of subjects experiencing an anaphylactic reaction by number of episodes, the number and percent of subjects experiencing an anaphylactic reaction by maximum severity using the Muraro grading scale (Muraro, 2007; Muraro, 2014), the number and percent of subjects experiencing an anaphylactic reaction that was a serious adverse event, the number and percent of subjects experiencing an anaphylactic reaction that required epinephrine use, the location of epinephrine episodes (home or study site), and the number of subjects experiencing an anaphylactic reaction that involved individual symptoms will be summarized for each study period.

### 9.8.4 Epinephrine Use as Rescue Medication

Epinephrine use is defined as any medication with a preferred name of 'EPINEPHRINE' when coded as described in Section 9.6.

All subjects per protocol are required to have epinephrine auto-injectors for use in case of a suspected anaphylactic reaction occurring outside the study site. There are differences in how and whether physicians record the prescription of epinephrine auto-injectors for PRN use. As a result, the presence or absence of a PRN prescription for epinephrine cannot be taken to indicate epinephrine usage, regardless of whether the prescription was written before or after study enrollment. The importance is counting the number of subjects receiving doses of epinephrine and the number of doses. As epinephrine should only be administered to treat a discrete allergic reaction, each dose of epinephrine should be closely temporally associated with a specific safety event and its use recorded as rescue medication on the concomitant medication case report form. In cases where a PRN epinephrine prescription is issued after the start of study product dosing, the study sites will be queried regarding if and when epinephrine was administered for a specific event.

The number and percent of subjects with at least 1 episode, where episode refers to 1 or more doses of epinephrine within a 2-hour window, will be summarized as well as the number of episodes experienced by each subject (1, 2, 3, and > 3). The total number of episodes will also be presented along with the number and percent of episodes by number of doses per episode, by associated adverse event severity, by seriousness of associated adverse events, by relatedness of associated adverse events, and by location of episode (home/study site). For subject counts, the number of subjects at risk within each study period will be used as the denominators. For unique episode counts, the total number of episodes within each study period will be used as the denominators. Data will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), study period (initial escalation, up-dosing, 300 mg QD, overall), and treatment group (AR101, placebo).

### 9.8.5 Food Allergy Episodes

Food allergy episodes will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), and study period (initial escalation, up-dosing, and 300 mg QD, overall). For each period, the number and percent of subjects experiencing any food allergy episode, the number and percent of subjects experiencing a food allergy episode in response to peanut (or nonpeanut), the number and percent of subjects experiencing any food allergy episode that was a serious adverse event, the number and percent of subjects experiencing any food allergy episode that required treatment, and the total number of food allergy episodes (peanut and nonpeanut related) will be summarized. Further, subject incidence of food allergy, peanut-related food allergy, and nonpeanut-related food allergy will be displayed by the number of episodes experienced.

All reported food allergy episodes will be listed by age group, treatment group, and subject. Allergic reactions associated with foods other than peanut will be flagged.

### 9.8.6 Allergy Symptoms During Dosing at the Study Site

During study product dosing at the study site, the severity of prespecified allergy symptoms is rated as mild, moderate, severe, life-threatening, or death. In addition, the presence of dose-limiting allergy symptoms is recorded for each dose.

The number of subjects experiencing any dose-related allergy symptoms during study product dosing at the study site and the maximum severity of symptoms will be summarized at each dose level. If a subject is administered the same dose at more than 1 study site visit (eg, dose was the same as the previous visit or dose was increased with a subsequent dose reduction), the most severe symptoms will be summarized for that dose level.

The time from dose administration to the time of onset of the first allergy symptom will be summarized using descriptive statistics and presented by dose level. Time from onset of the first symptom to resolution of the last symptom will be summarized similarly. If a subject receives more than 1 dose of study product at the study site at the same dose level, the subject will be counted only once using the shortest time from dose administration to onset of

the first symptom and the longest time from onset of first symptom to resolution of last symptom.

All summaries will be presented for the safety population by age group (4-11 years, 12-17 years, 4-17 years), and treatment group (AR101, placebo).

### 9.8.7 Allergy Symptoms During Dosing at Home

The allergic adverse event case report form contains all dose-related symptoms occurring at home. The number of subjects experiencing any dose-related symptoms during study product dosing at home and the maximum severity of symptoms will be summarized at each dose level. Summaries will be presented for the safety population by age group (4-11 years, 12-17 years, 4-17 years), and treatment group (AR101, placebo).

### 9.8.8 Assessment of Asthma Control

Asthma control in subjects with asthma was assessed using the ACT at baseline, up-dosing interim visit, end of up-dosing visit, early discontinuation visit, and exit visit in ARC007.

For subjects aged 12 years or older, the ACT has 5 questions, each recorded on a scale of 1 (least control) to 5 (greatest control). The total ACT is the sum of the 5 scores and ranges from 5 (least control) to 25 (greatest control). A total score of 19 or less indicates asthma is not adequately controlled. Missing data will not be imputed. If any of the 5 questions have a missing response, the total ACT score will not be calculated.

For subjects under aged 12 years, there are 4 questions for the subject and 3 questions for the parent/caregiver to complete. Subject responses range from 0 (least control) to 3 (greatest control). Parent/caregiver responses range from 0 (every day) to 5 (not at all). The sum of all 7 questions will make up the total score. The total ACT score for subjects under 12 years will range from 0 (least control) to 27 (greatest control). Missing data will not be imputed. If any of the questions have a missing response, the total ACT score will not be calculated for that subject.

All analyses of the ACT will be performed for the safety population by age group (4-11 years, 12-17 years) and treatment group (AR101, placebo). Summary statistics of the score for question, total score, and change from baseline will be tabulated by visit and treatment. A shift table of asthma control (adequate, not adequate, missing) will be summarized by treatment at each visit. The number of subjects with completed ACT questionnaires will be used as the denominator for all percentages.

Listing of the results from the questionnaire, including the total score, will be provided, sorted by age group, treatment group, subject, and visit.

### 9.8.9 Peanut-Specific IgE and IgG4

Blood samples to measure ps-IgE, ps-IgG4, and total IgE levels are collected during screening and at the study site visit where the 300 mg dose is administered in ARC007. Total IgE, ps-IgE, ps-IgG4, and calculated ps-IgE/IgG4 ratio will be listed by subject, and summarized by visit and treatment group. Results outside the limits of quantification will be displayed as less than the lower limit of quantification (LLOQ), or greater than the upper

limit of quantification (ULOQ), as appropriate. These values will be summarized as either the LLOQ or the ULOQ. If the ps-IgE or ps-IgG4 is outside of the limits of quantification, the ps-IgE/IgG4 ratio will be calculated using the LLOQ or ULOQ, as appropriate.

Summary statistics for the absolute values and change from baseline, including geometric means and geometric standard deviations, will be presented for the safety population by age group (4-11 years, 12-17 years, 4-17 years) and treatment group (AR101, placebo). Change from baseline ps-IgE, ps-IgG4, and ps-IgE/IgG4 ratio will be calculated on the $\log_{10}$ scale.

### 9.8.10 Skin Prick Test

The SPT is performed during screening, at the study site visit where the 300 mg dose is administered, and at the exit visit in ARC007.

Results from the SPT will be listed, including SPT device brand and type, test date, time, and measurements of the mean wheal diameter (in mm) of the following: peanut wheal (long axis), peanut wheal (short axis), saline wheal (long axis), saline wheal (short axis), histamine wheal (long axis), and histamine wheal (short axis). If environmental skin prick tests were performed, data from these tests will also be listed.

A derived mean wheal diameter score will be calculated as the average of the long and short axis from the peanut wheal minus the average of the long and short axis from the saline wheal. Summary statistics for the derived SPT mean wheal diameter and change from baseline will be presented for the safety population by age group (4-11 years, 12-17 years, 4-17-years) and treatment group (AR101, placebo).

### 9.8.11 Total Nasal Symptom Score

Assessment of nasal symptoms in subjects with allergic rhinitis using the TNSS questionnaire is performed at baseline, initial escalation, every 2 weeks during up-dosing, the early discontinuation visit, and the exit visit in ARC007.

The TNSS has 5 questions, where each question is assessed at 2 different timepoints (last 12 hours and last 2 weeks). The scores are recorded on a scale of 0 (none) to 3 (severe). The total score for each time point is the sum of the 5 scores and ranges from 0 (none) to 15 (severe). Missing data will not be imputed. If any of the 5 questions have a missing response, the total score will not be calculated.

All analyses for TNSS will be performed for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), visit, and time point. Summary statistics of the score for question, total score, and change from baseline will be tabulated.

Results from the questionnaire, including the total score, will be listed and sorted by age group, treatment group, subject, visit, and time point.

### 9.8.12 Quality of Life Assessments (FAQLQ and FAIM)

Quality of life assessment using the FAQLQ and the FAIM will be performed at the baseline and exit visit in ARC007.

Separate FAQLQ and FAIM instruments are administered based on age group. Parent/caregiver versions are also administered for subjects who are aged 17 years or younger. Due to differences between the various instruments, separate summaries will be provided by age group and person who completed the questionnaire (subject or parent/caregiver).

FAQLQ

The FAQLQ is a self-report instrument intended to assess the effect of food allergy on the subject's quality of life. Evaluations are done by the subject using a different form by age group (8-12 years and 13-17 years) (Flokstra-de Blok, 2008; Flokstra-de Blok 2009). Evaluations are also done by the parent/caregiver using a different form by age group (4-6 years, 7-12 years, and 13-17 years) (DunnGalvin, 2008).

Each question is scored from 1 to 7. The number of items and domains varies by instrument administered. For reporting, the domains for each included form are as follows:

Teen form –
a) Allergen Avoidance and Dietary Restrictions
b) Risk of Accidental Exposure
c) Emotional Impact

Child form –
a) Allergen Avoidance and Dietary Restrictions (Allergen Avoidance + Dietary Restrictions)
b) Risk of Accidental Exposure
c) Emotional Impact

Parent form teen –
a) Social and Dietary Limitations (Dietary Restrictions + Social Restrictions)
b) Food Anxiety
c) Emotional Impact

Parent form 7-12 –
a) Social and Dietary Limitations
b) Food Anxiety
c) Emotional Impact

Parent form 4-6 –
a) Social and Dietary Limitations
b) Food Anxiety
c) Emotional Impact

For each domain, the domain average score is the arithmetic average of the non-missing items comprising the domain. For all forms, the total score will be calculated as the average of the domain averages: (average a + average b + average c) / 3. For FAQLQ-PF 4-6, if items are completed that were not included for this age group, these items will not be included towards the scoring.

Descriptive statistics and scores of total and domain scores along with their changes from baseline will be provided. Data will be summarized separately by age group and responder (subject or parent/caregiver).

Listing of the raw scores as recorded in the case report form will be provided, sorted by treatment group and subject identification (ID).

FAIM

The FAIM is a self-report instrument intended to reflect the perception of food allergy severity and related risk as evaluated by the subject using a different form by age group (8-12 years and 13-17 years) (van der Velde, 2010). Evaluations are also done by the parent/caregiver using a different form by age group (4-12 years and 13-17 years). The instrument consists of 6 questions (4 expectation of outcome questions and 2 disease severity questions). The parent/caregiver versions include questions related to perception of disease severity and expectation of allergen exposure outcome. The parent/caregiver form includes 8 questions for subjects aged 4-12 years and 4 questions for subjects aged 13 to 17 years. Each question contains response options from 0 (limited severity perception) to 6 (greatest severity perception). The FAIM is scored on a 7-point scale from 1 to 7, and the total FAIM score is calculated as the arithmetic average of all nonmissing items.

Descriptive statistics for each item and the total scores along with their changes from baseline will be tabulated. Data will be summarized separately by age group and responder (subject or parent/caregiver).

Listing of the raw scores as recorded in the case report form will be provided, sorted by treatment group and subject ID.

### 9.8.13 Assessment of GI Symptoms by Pediatric Eosinophilic Esophagitis Symptom Scores

Subjects who discontinue treatment due wholly or in part to GI adverse events will be instructed to complete the Pediatric Eosinophilic Esophagitis Symptom Scores (PEESS v2.0) questionnaire (Martin, 2015; Franciosi, 2011) monthly for 6 months.

The PEESS questionnaire is composed of 20 items investigating the frequency and severity of eosinophilic esophagitis (EoE) symptoms in the last month. The total score consists of all 20 items. The frequency of symptoms is assessed by items 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, and 20, where each item is scored as: 0 = Never, 1 = Almost never, 2 = Sometimes, 3 = Often, 4 = Almost always. The severity of symptoms is assessed by items 2, 4, 6, 8, 10, 12, 14, 16, and 18, where each item is scored as 0 = Not bad at all, 1 = A little bad, 2 = Kind of bad, 3 = Bad, 4 = Very bad. Each item score is transformed to 0-100 as follows: 0 = 0, 1 = 25, 2 = 50, 3 = 75, 4 = 100.

The total, frequency total, and severity total scores are computed as the sum of the items divided by the number of items answered. If more than 50% of the items for the calculation of a score are missing, the score will not be calculated.

Summary statistics for frequency of symptoms, severity of symptoms, and the total score will be summarized for the safety population by age group (4-11 years on parent report only, 8-11 years on children and teen report only, 8-17 years on children and teen report only, 12-17 years, 4-17 years on parent report only), treatment group (AR101, placebo) and time point. PEESS results including the frequency total, severity total, and total scores will be listed.

### 9.8.14 Physical Examination

Physical examination results will be listed by age group, treatment group, subject, and visit.

### 9.8.15 Peak Expiratory Flow Rate and Spirometry

Peak expiratory flow rate (PEFR) assessments are performed throughout the study. Three attempts of PEFR are performed and the best (highest) value flagged in data listings. Only the best PEFR value will be summarized.

Spirometry may be performed at screening. Three attempts of $FEV_1$ are performed. Spirometry results are listed but are not summarized.

Observed values for PEFR as well as changes from baseline will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), and at each applicable visit. Results will be listed by age group, treatment group, subject, and visit.

### 9.8.16 Pregnancy Test Results

Pregnancy test results will be listed by age group, treatment group, subject, and visit.

### 9.8.17 Laboratory Assessments

Laboratory data consists of hematology tests. Laboratory data are collected at baseline and the 300 mg study site visit.

Observed values for each laboratory parameter as well as changes from baseline will be summarized for the safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), and at each applicable visit. For each laboratory parameter, a shift table based on the normal range (low, normal, and high) will be provided to summarize the baseline result versus the worst postbaseline result.

Results will be listed by age group, treatment group, subject, and visit.

### 9.8.18 Vital Signs

Vital signs (body temperature, heart rate, systolic/diastolic blood pressure) will be listed by subject and visit. Observed values and change from baseline will be summarized for the

safety population by age group (4-11 years, 12-17 years, 4-17 years), treatment group (AR101, placebo), and at each scheduled visit and time point. At initial escalation, up-dosing, and 300 mg QD visits, vital signs are taken predose and within 15 to 30 minutes after each dose is given. Vital signs are also taken at Exit and Early Discontinuation visits.

Additional vital signs measurements taken due to extension of the observation period will not be included in the summaries. All results will be listed by age group, treatment group, subject, visit, and time point.

## 10 PRE-DATABASE LOCK BLINDED DATA REVIEWS

Major data queries that have or may have bearing on the safety aspects of the study will be resolved prior to unblinding individual subjects. These include:

- Fulfillment of inclusion/exclusion criteria
- Accidental peanut exposures that resulted in serious adverse events or adverse events
- Severity and relationship to investigational product of serious adverse events and adverse events

The purpose of the predatabase lock blinded data review is for identification of major and minor protocol deviations.

Only ARC007 project team members who are blinded to study treatment assignments (ie, medical monitor, statistician, data manager, sponsor clinical project staff) will be involved in the data reviews.

## 11 REFERENCES

DunnGalvin A, de BlokFlokstra BM, Burks AW, Dubois AE, Hourihane JO. Food allergy QoL questionnaire for children aged 0-12 years: content, construct, and cross-cultural validity. Clin Exp Allergy. 2008;38(6):977-86.

Flokstra-de Blok BM, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JN, Duiverman EJ, Hourihane JO, et al. Development and validation of the self administered Food Allergy Quality of Life Questionnaire for adolescents. J Allergy Clin Immunol. 2008;122(1):139 44.

Flokstra-de Blok BM, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JN, Duiverman EJ, Hourihane JO, et al. Development and validation of a self administered Food Allergy Quality of Life Questionnaire for children. Clin Exp Allergy. 2009;39(1):127 37.

Franciosi JP, Hommel KA, DeBrosse CW, Greenberg AB, Greenler AJ, Abonia JP, et al. Development of a validated patient-reported symptom metric for pediatric eosinophilic esophagitis: qualitative methods. BMC Gastroenterol. 2011;11:126.

Martin LJ, Franciosi JP, Collins MH, Abonia JP, Lee JJ, Hommel KA, et al. Pediatric Eosinophilic Esophagitis Symptom Scores (PEESS v2.0) identify histologic and molecular correlates of the key clinical features of disease. J Allergy Clin Immunol. 2015;135(6):1519-28.

Muraro A, Roberts G, Clark A, Eigenmann PA, Halken S, Lack G, et al. The management of anaphylaxis in childhood: position paper of the European academy of allergology and clinical immunology. Allergy. 2007;62(8):857-71.

Muraro A, Roberts G, Worm M, Bilò MB, Brockow K, Fernández Rivas M, et al. Anaphylaxis: guidelines from the European Academy of Allergy and Clinical Immunology. Allergy. 2014;69(8):1026-45.

van der Velde JL, Flokstra-de Blok BM, Vlieg-Boerstra BJ, Oude Elberink JN, DunnGalvin A, Hourihane JO, et al. Development, validity and reliability of the food allergy independent measure (FAIM). Allergy. 2010;65(5):630-5.