STATISTICAL ANALYSIS PLAN

ARC004
Peanut Allergy Oral Immunotherapy Study of AR101 for Desensitization in Children and Adults (PALISADE) Follow-On Study

Version 1.0 – 10 Jul 2019
Reference Numbers: NCT02993107, EudraCT 2016-004941-94

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

Sponsor Name:	**Aimmune Therapeutics, Inc.**
Protocol Number and Title:	**ARC004** **Peanut Allergy Oral Immunotherapy Study of AR101 for Desensitization in Children and Adults (PALISADE) Follow-on Study** Amendment 5.0 (Global) 12 April 2018
Authors:	Lalith Akella Biostatistical Consultant Aimmune Therapeutics, Inc. Brian Beus / Laura Holland Precision for Medicine, Oncology and Rare Disease 6005 Hidden Valley Road, Suite 170 Carlsbad, CA 92011
SAP Version:	Version 1.0
SAP Version Date:	10Jul2019

Version: **Version 1.0** **Version Date:** **10JUL2019**
Document Effective Date:
Document File Name:

I confirm that I have reviewed this document and agree with the content.

Precision for Medicine, Oncology and Rare Disease

Laura Holland, MS
Biostatistician II — 12-Jul-2019 — Date

Brian Beus, MS
Principal Biostatistician — 12 Jul 2019 — Date

Aimmune Therapeutics, Inc

Daniel Adelman, MD
Chief Medical Officer — 12 July 2019 — Date

Rezi Zawadzki, DrPH
Vice President, Biometrics — 12 July 2019 — Date

Lalith Akella
Biostatistical Consultant, Biometrics — 10 July 2019 — Date

TABLE OF CONTENTS

GLOSSARY OF ABBREVIATIONS

Abbreviation	Description
ACT	Asthma Control Test
AE	Adverse Event
AR101	Characterized Peanut Allergen
ATC	Anatomical Therapeutic Chemical
BIW	Twice Weekly
CI	Confidence Interval
CoFAR	Consortium of Food Allergy Research
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
DBPCFC	Double-Blind, Placebo-Controlled Food Challenge
EM	Extended Maintenance
FAIM	Food Allergy Independent Measure
FAQLQ	Food Allergy Related Quality of Life Questionnaire
FEV_1	Forced Expiratory Volume in one second
FVC	Forced Vital Capacity
GI	Gastrointestinal
ICH	International Conference on Harmonization
IgE	Immunoglobulin E
IgG	Immunoglobulin G
LLOQ	Lower Limit of Quantification
MedDRA	Medical Dictionary for Regulatory Activities
MTD	Maximum Tolerated Dose
OIT	Oral Immunotherapy
PEF	Peak Expiratory Flow
PRACTALL	Practical Allergy, a joint initiative between the European Academy of Allergy and Immunology and The American Academy of Asthma, Allergy and Immunology
PRN	As needed (pro re nata)

Abbreviation	Description
ps	Peanut-specific
PT	Preferred Term
QD	Once Daily
QOD	Every Other Day
QoL	Quality of Life
QOW	Every Other Week
QW	Once Weekly
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SOC	System Organ Class
SPT	Skin Prick Test
TEAE	Treatment-Emergent Adverse Event
TLF	Table, Listing and Figure
TSQM-9	Treatment Satisfaction Questionnaire for Medication
ULOQ	Upper Limit of Quantification
WHO-DDE	World Health Organization Drug Dictionary

1. PURPOSE

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

1.1. DOCUMENT HISTORY

This is the first version.

1.2. RESPONSIBILITIES

The SAP was prepared by Precision for Medicine, Oncology and Rare Disease ("Precision"). Precision will perform the final statistical analyses and be responsible for the production and quality control of all tables, listings and figures (TLFs).

1.3. TIMING OF ANALYSES

Interim Analysis No interim analysis of efficacy is planned for this study.

Final Analysis The final analysis of safety and efficacy is planned after all subjects complete Exit/Early Discontinuation Visit assessments. The final analysis will include all data collected through to the time of database lock.

2. STUDY OBJECTIVES

2.1. PRIMARY OBJECTIVE

The primary objective of this study is to determine the safety, tolerability, and efficacy of AR101 using a characterized oral immunotherapy (OIT) desensitization approach and alternative maintenance dosing intervals.

2.2. SECONDARY OBJECTIVES

The secondary objectives of the study are:

- To confirm the safety profile of AR101 as measured by the incidence of adverse events (AEs), including serious adverse events (SAEs).
- To confirm the efficacy of AR101 through reduction in clinical reactivity, measured in a Double-Blind, Placebo-Controlled Food Challenge (DBPCFC) to a cumulative dose of 4043 mg.
- To evaluate subjects' quality of life (QoL) and treatment satisfaction during AR101 treatment on daily and nondaily treatment regimens.
- To evaluate the long-term immunologic effects of AR101 treatment.

2.3. BRIEF DESCRIPTION OF STUDY DESIGN

This is an international, multicenter, open-label, 2-arm follow-on study of ARC003. The 2 arms to be studied in ARC004 are the subjects who received placebo in ARC003 (Group 1) and those who received AR101 in ARC003 (Group 2).

Group 1 (former placebo): Subjects who complete the placebo arm of ARC003 are eligible to enroll in ARC004. All former placebo-treated subjects who consent to participate in ARC004 will enter Group 1 and will undergo initial dose escalation, up-dosing, and maintenance with AR101, as in the AR101 arm of ARC003. The Initial Dose Escalation Period consists of a step-wise dose-escalation from 0.5 to 3 or 6 mg (as tolerated) on Day 1 and confirmation of the ability to tolerate a single dose of 3 mg on Day 2. The Up-dosing Period is 22 to 40 weeks in duration and consists of dose escalations every 2 weeks up to a maximum of 300 mg/d (as tolerated). The Maintenance Period consists of daily dosing at 300 mg/d for approximately 24 weeks, at which point Group 1 subjects will undergo a DBPCFC to a maximum total of 4043 mg cumulative peanut protein to test the efficacy of AR101 in desensitizing ARC003 former placebo-treated subjects to large amounts of peanut protein after 6 months of daily maintenance.

All Group 1 subjects tolerating ≥ 443 mg of cumulative peanut protein in this challenge will be eligible to continue in ARC004, and enter the Extended Maintenance (EM) Period, initially consisting of ongoing daily maintenance therapy at 300 mg/d. Dependent upon the

EM results from ARC004 Group 2 (below), subjects in Group 1 EM will test the gradual lengthening of their dosing interval from once daily (QD) to every other day (QOD), twice weekly (BIW), once weekly (QW), and finally every other week (QOW), as tolerated. The length of each interval will be between 8 and 24 weeks and will be adjusted based on the data from Group 2. Lengthening to each level will occur sequentially (QD then QOD then BIW then QW and finally QOW), and only if sufficient evidence exists, based on the analyses of the Group 2 EM subjects and the ongoing safety of Group 1 subjects, to support progression to each level.

Group 2 (former AR101): Subjects who complete the AR101 arm of ARC003 and tolerate a challenge dose of ≥ 443 mg cumulative of peanut protein at the ARC003 Exit DBPCFC are eligible to enroll in ARC004. Group 2 subjects will have undergone the Initial Dose Escalation, Up-dosing, and Maintenance Periods in ARC003 and will therefore enter the ARC004 EM period directly.

Upon entry into the EM phase of the study, Group 2 subjects will be consecutively enrolled into 1 of several cohorts that will conditionally explore an alternate dosing interval regimen over 28-week study periods, illustrated in Figure 1 and as follows:

EM Cohort 1: The first 120 Group 2 subjects to enter ARC004 will comprise Cohort 1, and will continue 300 mg of AR101 QD for 28 weeks before undergoing an Exit DBPCFC (total 28 weeks).

EM Cohort 2: The next 50 Group 2 subjects enrolling in ARC004 will comprise Cohort 2 and will take 300 mg of AR101 QOD for 4 weeks and then BIW (eg, Monday/Thursday) for 24 weeks (total of 28 weeks) as tolerated, before undergoing an Exit DBPCFC.

EM Cohort 3: All remaining subjects recruited into ARC004 will comprise Cohort 3. This cohort will be randomized 1:1:1 to 1 of 3 initial strategies:

a. 300 mg QD for 56 weeks followed by an Exit DBPCFC (Cohort 3A) (total 56 weeks)
b. 300 mg QD for 28 weeks, then 300 mg QOD for 4 weeks, then BIW for 24 weeks, as tolerated, followed by an Exit DBPCFC (Cohort 3B) (total 56 weeks)
c. 300 mg QD for 28 weeks, then 300 mg QOD for 4 weeks, then BIW for 24 weeks, then 300 mg QW for 28 weeks, as tolerated, followed by an Exit DBPCFC (Cohort 3C) (total 84 weeks)

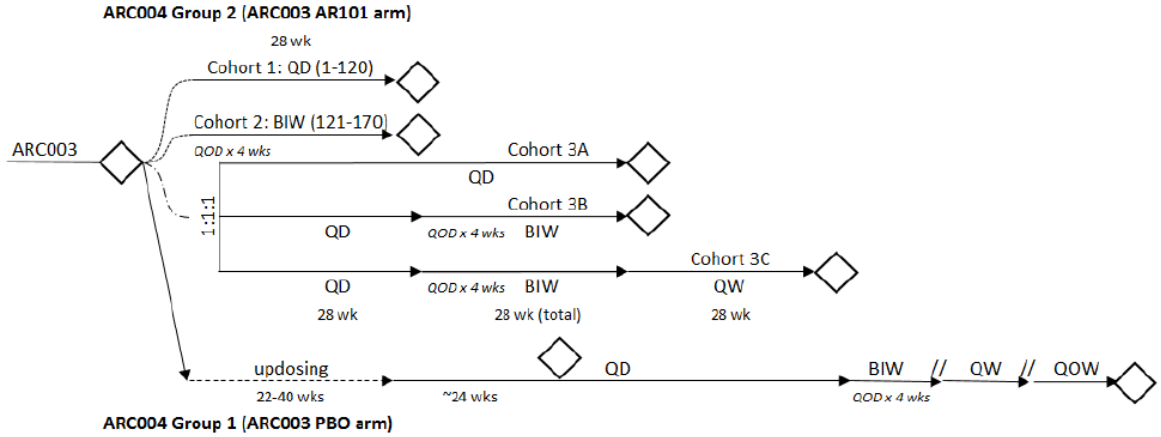
In ARC004, Exit DBPCFCs will test to a maximum of 4043 mg cumulative peanut protein to determine the continued efficacy of each alternate dosing schedule. Following the end-of-study visit, each subject's participation in ARC004 ends. The end of the study is defined as the date of the last visit of the last subject.

As per the study design, if subjects were on any other regimen other than QD and were tolerating it well during ARC003, those subjects were allowed to continue on their regimen in ARC004. As per the safety monitoring committee suggestion, all subjects in group 2

cohort 3C who were on BIW, should remain on BIW. All group 1 subjects who were on QD should remain on QD throughout the remainder of the study. No specific safety signals were identified but overall data did not support increasing the interval of dosing. If subjects were tolerating BIW then they may continue with that frequency.

Figure 1. Illustration of study design

Figure 1: Study Schematic



Diamonds represent DBPCFCs; all challenges in ARC004 will be to a total of 4043 mg cumulative peanut protein, according to the table above.
The initial ARC004 DBPCFC for Group 1 will be completed after approximately 24-week maintenance period
QD = daily, QOD = every other day, BIW = twice weekly, QW = once weekly, QOW = every other week
NOTE: The BIW interval begins with 4 weeks of QOD dosing before transitioning to 24 weeks of BIW dosing, as tolerated
Group 1: Total duration varies, ranging from approximately 88 to 136 weeks
Group 2, Cohort 1: 28 weeks
Group 2, Cohort 2: 28 weeks
Group 2, Cohort 3A: 56 weeks
Group 2, Cohort 3B: 56 weeks
Group 2, Cohort 3C: 84 weeks

2.4. DETERMINATION OF SAMPLE SIZE

As stated in the protocol, the sample size for the present study is based on the number of eligible subjects from study ARC003 consenting to rollover into ARC004:

"ARC003 had a target enrollment of 500 and a 3:1 AR101: placebo randomization ratio, providing 89% power to rule out a treatment difference (AR101 minus placebo) of 15% or less at the primary endpoint of the proportion of subjects who tolerate ≥ 1043 mg cumulative of peanut protein with no more than mild symptoms at the Exit DBPCFC. Assuming accrual of 80% of the total ARC003 sample into ARC004, the total of approximately 400 subjects treated with AR101 in the clinical program would provide an 80% probability of observing ≥ 1 AE when the background rate of the AE is 4 per 1,000 subjects. There is no prospective power calculation for efficacy in ARC004."

3. ENDPOINTS

3.1. PRIMARY ENDPOINT

The primary endpoint is the frequency of treatment-related AEs, including SAEs, during the overall study period (from enrollment to discontinuation/exit). ARC004 is specifically designed to begin exploring alternate (eg, nondaily) oral immunotherapy (OIT) extended maintenance dosing regimens. The data generated for alternative dosing regimens in ARC004 are expected to be exploratory and hypothesis-generating.

3.2. SECONDARY ENDPOINTS

The secondary safety endpoints are as follows:

- Frequency of anaphylactic reaction
- Frequency of use of epinephrine as a rescue medication
- Frequency of AEs leading to withdrawal of AR101
- Frequency of AEs in each treatment regimen leading to discontinuation of extended interval dosing
- Frequency of Gastrointestinal (GI) AEIs
- Occurrence of premature discontinuation of dosing due to chronic/recurrent gastrointestinal (GI) AEs
- Frequency of accidental food allergen exposure
- In subjects with asthma, change in asthma control using the Asthma Control Test (ACT) questionnaire
- Frequency of all above safety endpoints by treatment period

The secondary efficacy endpoints are as follows:

- The proportion of subjects in each regimen tolerating a single dose of 600 mg (≥ 1043 mg cumulative) of peanut protein during their Exit DBPCFC
- The proportion of subjects in each regimen who tolerate a single dose of 300 mg (≥ 443 mg cumulative) of peanut protein during their Exit DBPCFC
- The proportion of subjects in each regimen who tolerate a single dose of 2000 mg (4043 mg cumulative) of peanut protein during their Exit DBPCFC
- Maximum tolerated dose and change from baseline* at the Maintenance DBPCFC** and each Exit DBPCFC
- Maximum severity of symptoms at each challenge dose at the Maintenance DBPCFC** and each Exit DBPCFC

- Frequency of use of epinephrine as a rescue medication at the Maintenance DBPCFC** and each Exit DBPCFC
- Change in QoL as assessed by the food allergy related quality of life questionnaire (FAQLQ) and the food allergy independent measure (FAIM) questionnaires
- Satisfaction with AR101 treatment as assessed by the TSMQ-9 questionnaire and additional questions
- Changes in peanut-specific (ps) serum immunoglobulin E (IgE) and IgG4 levels
- Changes in peanut skin prick test (SPT) wheal diameter

*Baseline is defined as the Exit DBPCFC from ARC003 for both treatment groups.

**Maintenance DBPCFC is defined only for Group 1.

4. ANALYSIS POPULATIONS

The following analysis populations will be defined for this study.

4.1. SAFETY POPULATION

The primary population of analysis for all analyses will be the Safety population. The Safety population will consist of all subjects who receive at least one dose of study treatment during ARC004.

4.2. COMPLETER POPULATION

The Completer population will include all subjects in the safety population who have at least one evaluable maintenance or exit DBPCFC, where an evaluable DBPCFC is defined as completion of at least the peanut part of the food challenge.

5. PROTOCOL DEVIATIONS

All protocol deviations will be reported on a specific case report form (CRF) as follows:

- Inclusion Criteria
- Exclusion Criteria
- Received incorrect study treatment
- Randomization issue/ randomized to wrong stratum
- ICF
- SAE Not Reported
- Visit Out of Window
- Missed Study Visit
- Procedure Not Per Protocol
- Prohibited Concomitant Medication
- Lab Sample missed
- Study Drug Compliance
- Other (with free text field to record detail)

Protocol deviations will be reviewed and categorized into major or minor prior to database lock. If the PP population is defined, the major protocol deviations causing exclusion from the PP population will be identified. All major protocol deviations will be listed and included in the study report.

6. GENERAL ASPECTS FOR STATISTICAL ANALYSIS

6.1. STATISTICAL NOTATION AND METHODOLOGY

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 non-missing 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". Select ordinal data may be summarized using both descriptive statistics and counts and percentages of subjects in each category, as appropriate.

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. Other statistics (e.g., confidence intervals (CIs)) will be presented using the same general rules outlined above, or assessed for the most appropriate presentation based on the underlying data.

Unless otherwise noted, values reported as greater than or less than some quantifiable limit (e.g., "< 1.0") will be summarized with the sign suppressed in summary tables and figures, using the numeric value reported. Data will display on subject listings to include the sign.

Summary tables will be presented by group/cohort displayed as Group 1 and Group 2 with Cohort 1, Cohort 2, Cohort 3A, Cohort 3B, and Cohort 3C. For disposition, demographic, and other summaries of baseline and history data, a Total column for both groups combined will be included. For any summaries of events such as adverse events or concomitant medications, event incidences will be summarized in the Safety population separately for the following group/cohort study periods:

- Group 1: IDE and Up-Dosing, all QD Maintenance (including extended maintenance), Overall
- Group 2 Cohort 2: QOD, BIW, and Overall
- Group 2 Cohort 1 (includes QD only)
- Group 2 Cohort 3A (includes QD only)
- Group 2 combined cohorts B and C: QD, QOD, BIW, and Overall
- Group 2 Overall QD

For subjects in Group 2 who were not able to tolerate BIW or QOD dosing per study design and who subsequently reverted to a QD regimen, safety events for these subjects while on QD will be flagged in listings. Safety data collected after the dosing change will not be summarized in tables but will be presented in separate listings for key safety endpoints.

No summaries for any QW or QOW dosing regimens will be tabulated since no subjects initiated these dosing regimens.

All relevant data collected in the database and any derived data will be included in data listings and unless otherwise noted sorted by age group, group/cohort, subject number, test/measurement, and visit and time point as appropriate.

All statistical methods will be based on the International Conference on Harmonization (ICH)-E9 Guidance for Industry "Statistical Principles for Clinical Trials".

A review of the database will be conducted before database lock. Any decision to amend the planned statistical analysis will be documented in an amendment to the statistical analysis plan and details will be included in the CSR.

If the assumptions underlying the formal statistical methods proposed are not met during the analysis of the final data, an alternative, more appropriate, statistical method will be used and any changes documented in the statistical methods section of the CSR, including the rationale for use.

Additional exploratory analyses of the data will be conducted as deemed appropriate. These analyses will be fully documented and clearly identified as post-hoc and exploratory.

6.2. STRATA AND COVARIATES

No statistical modeling will be performed for this study, and no stratification was applied to the group or cohort assignments in this open-label, follow-on study.

6.3. SUBGROUP ANALYSES

In general, summaries of the primary and secondary endpoints will be presented for the children aged 4-17 years. Additionally, key safety tables will be presented for 18-55 year old subjects.

6.4. MULTIPLE COMPARISONS AND MULTIPLICITY

No formal statistical comparisons will be performed in this study.

6.5. SIGNIFICANCE LEVEL

Unless stated otherwise, all CIs will be calculated at the 95% level and provided for descriptive purposes only rather than formal statistical inference.

7. DATA HANDLING METHODS

7.1. 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.

7.2. DATA PRESENTATION

Individual subject data listings will be provided to support summary tables and serve as a data source. Unless otherwise noted, all data collected during the study for all enrolled subjects will be included in data listings.

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 post-baseline values for laboratory results).

7.3. MAXIMUM TOLERATED DOSE AT DBPCFC

The maximum tolerated dose (MTD) for a DBPCFC is defined as the maximum single dose of peanut protein resulting in no more than mild symptoms and determined to have been tolerated (i.e., the subject did not experience any dose-limiting symptoms). Any symptom requiring treatment is inherently dose-limiting; thus, a dose during a DBPCFC cannot be considered “tolerated” if treatment was deemed necessary by the investigator. The MTD at the ARC003 Exit DBPCFC will be used as the baseline amount of peanut protein tolerated. If a subject is administered non-standard doses at a DBPCFC, the MTD will be considered as the highest standard dose (whether administered or not) that is less than the highest tolerated non-standard dose.

When describing the MTD at the Exit DBPCFC, in terms of the cumulative amount of peanut protein, the 1 mg dose will not be included. Thus, subjects who tolerate all dose levels from 3 mg to 300 mg or 1 mg to 300 mg have a cumulative MTD of 443 mg.

7.4. CLASSIFICATION OF RESPONDER STATUS

7.4.1. Classification as a responder

To be a 600 mg responder at the EXIT DBPCFC, a subject must meet both of the following conditions:

1) Must have attained an MTD ≥ 600 mg of peanut protein on Exit DBPCFC
2) Must not have experienced more than mild symptoms through 600 mg of peanut protein on Exit DBPCFC.

Classification as a 300 mg, a 2000 mg, and a 1000 mg responder are defined similarly.

A 600 mg responder is equivalent to tolerating ≥ 1043 mg cumulative of peanut protein during their Exit DBPCFC. Likewise a 443 mg cumulative corresponds to a 300 mg responder, a 2043 mg cumulative dose corresponds to 1000 mg responder, and 4043 mg cumulative corresponds to a 2000 mg responder.

Responders at maintenance DBPCFC for Group 1 subjects are defined similarly.

7.4.2. Classification as a non-responder

If a subject cannot be classified as a responder at a given dose level, then that subject should be classified as a non-responder at that dose level.

7.5. DATA DERIVATIONS AND DEFINITIONS

The following definitions and derivations will be used throughout this study:

- Study day is calculated relative to the first dose date in ARC004 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.
- In this follow on study, baseline is defined as the last non-missing value prior to the first dose of study treatment in the study where study product is first received. For Group 1 patients (those previously on placebo in ARC003), baseline is relative to first dose of AR101 in study ARC004. For Group 2 patients (those continuing on AR101 therapy from ARC003), baseline is relative to first dose of AR101 in study ARC003. If no ARC004 data were entered for a Group 1 subject, use the last available ARC003 data prior to first dose of AR101 in ARC004. For food challenge end points, the ARC003 exit food challenge will serve as the baseline for both groups.

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

- Group 1 Initial Dose Escalation period is defined as the time period beginning with the date and time of the first dose of study product in clinic and ending with the date of the last dose of study product taken prior to Up-dosing.

- Group 1 Up-dosing period is defined as the time period beginning with the date and time of the first home dose of study product at 3 mg, and ending with the date and time of first in-clinic dose at 300 mg. This period will be, ideally, 22 weeks in duration, but may be extended to a maximum of 40 weeks to accommodate dose reductions and re-escalations, if necessary.

- Group 1 Maintenance period is defined as the time period beginning with the date and time of the first home dose of study product at 300 mg and ending with the date of the last dose of study product taken prior to the maintenance DBPCFC. The maintenance period lasts approximately 24 weeks.

- Extended Maintenance (EM) period for Group 1 is defined as the time period beginning with the date and time of first dose of study product after the maintenance DBPCFC and ending with the date of the last dose of study product taken prior to Exit DBPCFC. Group 2 subjects will have undergone the Initial Dose Escalation, Up-dosing, and Maintenance Periods in ARC003 and will therefore enter the ARC004 EM period directly. The EM period for Group 2 is defined as the time period beginning with the date and time of first dose of study product in ARC004 and ending with the date of the last dose of study product taken prior to Exit DBPCFC. Length of EM period vary by group and cohort. See section 2.3 for additional detail about each group and cohort.

- Exit DBPCFC period is defined as the time period beginning with the date and time of the first Exit DBPCFC dose and through 24 hours after the last dose of DBPCFC product.

- A DBPCFC is given in 2 parts, either on the same day, on consecutive days, or occasionally on non-consecutive days. For the non-consecutive days, the period of time more than 24 hours after the first part and before the second part begins will be attributed as follows

    1. For the Group 1 maintenance DBPCFC this is attributed to the Maintenance Period

    2. For the Exit DBPCFC Period it is attributed to the EM period.

- The active treatment period is defined as the time period beginning with the date and time of the first dose of study product in ARC004 and ending with the date and time of the last dose of study product in ACR004.

- Duration of active treatment period (days) is calculated as the date of last dose in ARC004 minus the date of first dose in ACR004 plus 1, excluding the DBPCFC periods.

- The Exit DBPCFC is defined as indeterminate if the subject was not able to tolerate the placebo challenge up to and including a dose of 2000 mg.

7.6. MISSING DATA

All AEs with partial/missing dates and times will be considered Treatment-Emergent Adverse Events (TEAEs) unless a partial date clearly indicates that it occurred prior to first dose of study treatment or more than 30 days after the last dose of treatment. All therapies with partial or missing dates and times recorded on the Concomitant Medication or Non-Drug Therapy CRF 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 which have 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 Day 1 date will be used
- In all other cases the 1st of the month will be used with the provided month and year

Start dates with a missing day and month but which have 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 the 1st of January 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 most recent reaction to peanut on the peanut allergy history CRF 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 Jan 1 if the same year as the informed consent date. Otherwise it is set to July 1.

For any anaphylactic reactions that are missing severity, severe, the highest severity on the Muraro scale, will be imputed. For any allergic reactions that are missing severity, severe, the highest rating on the Consortium of Food Allergy Research (CoFAR) scale, will be imputed. For any other adverse events that are missing severity, severe, the highest rating on the Common Terminology Criteria for Adverse Events (CTCAE) scale, will be imputed.

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

7.7. POOLING

In general for summaries of event data, Group 2 cohorts 3B and 3C will be combined and summarized together. For overall summaries of QD exposure in Group 2, the QD exposure period will be pooled across all Group 2 cohorts.

8. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS

8.1. SUBJECT DISPOSITION AND WITHDRAWALS

The number of subjects enrolled as well as the number and percentage of subjects who entered each study period, and were included in each analysis population will be summarized by age group (ages 4-17, 18-55) and group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, Group 2 Cohort 3C, and Total).

Primary reason for early discontinuation from the study will be summarized as well as whether subjects underwent the exit DBPCFC.

Inclusion and exclusion eligibility will be listed separately.

8.2. PROTOCOL DEVIATIONS

All protocol deviations as defined in Section 5 will be listed by subject. Major protocol deviations will be summarized for the Safety population for age group 4-17 years by group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, Group 2 Cohort 3C, and Total).

8.3. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Summary statistics for demographic and baseline characteristics will be provided for the Safety population summarized by age group (ages 4-17, 18-55) and group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, Group 2 Cohort 3C, and Total).

Demographic data will include age, sex, country, ethnicity, race, height, body weight, and body mass index (BMI). Baseline characteristics include total IgE, ps-IgE, ps-IgG4, ps-IgE/IgG4 ratio, SPT mean wheal diameter at screening, MTD of peanut protein at ARC003 Exit DBPCFC, childbearing potential, method of birth control, and history of asthma.

Demographic characteristics that cannot change over time (e.g., race) are summarized as collected in ARC003.

Age will be calculated relative to date of informed consent for the ARC004 study, 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 one;

- 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.

8.4. PEANUT ALLERGY HISTORY

The duration of peanut allergy (months since peanut allergy diagnosis), number of anaphylactic reactions to peanut in lifetime, months since most recent allergic reaction, and the symptoms experienced during the most recent peanut exposure will be summarized for the Safety population for age group 4 to 17 years by group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, Group 2 Cohort 3C, and Total).

For Group 2, peanut allergy history is taken as entered in the ARC003 CRF. For Group 1, peanut allergy history is taken as entered in the ARC004 CRF, except for date of first diagnosis of peanut allergy which is reported as collected in ARC003 because it does not change over time.

The reported date of the most recent allergic reaction and date of diagnosis of peanut allergy will be imputed based on the logic in section 7.6.

Peanut allergy history will be listed for all subjects.

8.5. NONPEANUT ALLERGY HISTORY

Nonpeanut allergy history will be listed for all subjects. The presence of nonpeanut allergy history and type of allergy will also be summarized for the safety population for age group 4 to 17 years by group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, Group 2 Cohort 3C, and Total).

For Group 2, nonpeanut allergy history will be taken as entered in the CRF in ARC003. For Group 1, nonpeanut allergy history will be taken as entered in the CRF in ARC004.

8.6. OTHER MEDICAL HISTORY

Medical history will be listed for all subjects. Subjects with abnormal medical history events will be summarized for the safety population for age group 4 to 17 years by group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, Group 2 Cohort 3C, and Total) and by the Medical Dictionary for Regulatory Activities (MedDRA) system organ class (SOC) version 19.1 and preferred term (PT).

Medical history will be considered prior to the start of study product dosing. For Group 2, medical history will be taken as entered in the CRF in ARC003. For Group 1, medical history will be taken as entered in the CRF in ARC004.

9. EFFICACY

9.1. ANALYSIS OF EFFICACY SECONDARY ENDPOINTS

To confirm the efficacy of AR101 through reduction in clinical reactivity as measured in a DBPCFC to a cumulative dose of 4043 mg, the Exit food challenge will be summarized by group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, and Group 2 Cohort 3C). For Group 1, the maintenance DBPCFC will be summarized separately. Non-food-challenge efficacy endpoints also will be summarized by group/cohort (Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, and Group 2 Cohort 3C). Except where otherwise noted, DBPCFC summaries will be included for the completer population for age group 4-17 years.

The study examines DBPCFC changes from the baseline of study ARC004 (i.e. DBPCFC completed at end of study ARC003) to the end of study ARC004.

9.1.1. Desensitization Response Rates

Desensitization response rate refers to the proportion of subjects who achieve desensitization as determined by tolerating the dose levels specified in section 7.4 of peanut protein with no more than mild symptoms at the DBPCFC. These subjects are termed responders.

All subjects failing to achieve the definition of responder are termed non-responders.

The desensitization response rates at Exit DBPCFC and their associated 95% confidence intervals using exact Clopper-Pearson confidence intervals will be presented for each group/cohort. Desensitization response rates at the maintenance DBPCFC will be presented similarly for Group 1 subjects who complete the maintenance DBPCFC.

The number and percent of subjects at each dose level for highest tolerated dose will also be summarized at the Exit DBPCFC for each group/cohort and at the maintenance DBPCFC for Group 1.

9.1.2. Maximum Tolerated Dose at Exit DBPCFC

The DBPCFCs are based on a modified Practical Allergy (PRACTALL) dosing regimen as described in the ARC003 protocol (Section 3). ARC004 DBPCFCs will include a single highest dose of 2000 mg, whereas ARC003 included a single highest dose of 1000 mg at Exit. With the exception of the 600 mg and 2000 mg doses, the modified PRACTALL doses are approximately on a logarithmic scale.

The MTD will be summarized at screening and exit along with fold increase from screening for each group/cohort. Subjects who did not tolerate any dose will be assigned an MTD of 0.3 mg to calculate fold change. Changes from screening to the maintenance DBPCFC will be summarized similarly for Group 1 for subjects who complete the maintenance DBPCFC.

9.1.3. Maximum Severity of Symptoms at DBPCFC

Maximum severity of symptoms at the DBPCFC will be summarized at the 2000 mg, 1000 mg, and 600 mg levels. Symptom severity is determined according to the CoFAR scale at 5 levels: 1-Mild, 2-Moderate, 3-Severe, 4-Life Threatening, and 5-Fatal. Analysis is done on 4 levels: 0-None, 1-Mild, 2-Moderate, 3-Severe or higher (severe, life threatening, fatal). Subjects who experience no symptoms will be assigned a severity of 0-None. Symptom severity data is collected at each challenge dose of peanut protein during the DBPCFC (3 mg, 10 mg, 30 mg, 100 mg, 300 mg, 600 mg, 1000 mg, and 2000 mg). The maximum severity of symptoms observed in the DBPCFC at the specified dose (2000 mg, 1000 mg, or 600 mg) or lower will be used for each subject.

The number and percent of subjects will be presented by maximum severity for the specified challenge dose levels at the Exit DBPCFC for each group/cohort and at the maintenance DBPCFC for Group 1 for subjects who complete the maintenance DBPCFC.

9.1.4. Use of Epinephrine as a Rescue Medication at DBPCFC

The number and percentage of subjects using Epinephrine as a rescue medication at the Exit DBPCFC will be summarized by type of challenge (peanut protein or placebo) for each group/cohort. Similarly, the number and percentage of subjects using Epinephrine as a rescue medication at the maintenance DBPCFC will be summarized by type of challenge (peanut or placebo) for Group 1.

The number and percentage of subjects using epinephrine use as a rescue medication will also be summarized by each dose level (up through 2000 mg, up through 1000 mg, up through 600 mg, up through 300 mg, etc.) at the Exit peanut DBPCFC for each group/cohort and at the maintenance peanut DBPCFC for Group 1.

The number of epinephrine doses as a rescue medication will be summarized by dose level (up through 2000 mg, up through 1000 mg, up through 600 mg, etc.) during the Exit peanut DBPCFC for each group/cohort and during the maintenance peanut DBPCFC for Group 1. Doses will be categorized as 0 doses, 1 dose, 2 doses, and 3 or more doses.

Listings of all Epinephrine use will be provided.

9.1.5. Changes in Peanut-specific IgE, total IgE and peanut-specific IgG_4

Blood samples will be collected to measure ps-IgE, ps-IgG4 levels, and total IgE. Ps-IgE/IgG4 ratio will be calculated, listed by subject, and summarized by visit and group/cohort. 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-IgG_4 is outside of the limits of quantification, the ps-IgE/IgG4 ratio will be calculated using the LLOQ or ULOQ as appropriate.

Summary statistics, including geometric means and geometric standard deviations, will be presented by visit and group/cohort, and change from baseline to each visit will also be presented.

9.1.6. Peanut Skin Prick Test

Results from the SPT will be listed, including test date and time, and measurements of the mean wheal diameter (in mm) of the following: peanut wheal (long axis), peanut erythema/flare (short axis), saline wheal (long axis), saline-glycerin erythema/flare (short axis), histamine wheal (long axis), and histamine erythema/flare (short axis).

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 will be presented at each visit by group/cohort and change from baseline to each visit will be presented.

9.1.7. Quality of Life Assessments

Quality of life assessment using the FAQLQ and the FAIM questionnaires will be performed prior to the start of treatment and after the Exit DBPCFC or early discontinuation.

Separate FAQLQ and FAIM instruments are administered based on the subject's age group. Parent versions are also administered for all subjects. Due to differences between the various instruments, separate summaries will be provided by age group and person who completed the questionnaire (i.e., subject or parent).

FAQLQ:

The FAQLQ is a self-report instrument that is 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 presented on a scale from 0 (not at all) to 6 (extremely) and 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:

Adult form ≥ 18 –

a) Allergen Avoidance and Dietary Restrictions
b) Risk of Accidental Exposure
c) Emotional Impact
d) Food Allergy Related Health

Teen form 13-17 –

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

Child form 8-12 –

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

Parent form teen 13-17 –

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 any respondent and form, the domain score algorithms above should be derived only if more than half the items comprising that domain score have valid responses. For all forms, the total score will be calculated as the average of the domain averages: (average a + average b + average c) / 3. Total scores should not be

computed if one or more of the domain scores comprising the total are missing. 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. For FAQLQ-AF, the fourth domain of Food Allergy Related Health is not included in the total score calculation.

For child, teen, and all parent forms, descriptive statistics on the total and domain scores along with their changes from baseline will be provided. Data will be summarized for the safety population separately by questionnaire age group and responder (subject or caregiver).

Listing of the raw scores as recorded in the CRF, domain scores, and total scores will be provided, sorted by age group, group/cohort, and subject ID.

FAIM:

The FAIM is a self-report instrument that is 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+ 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.

The FAIM is scored on a 7-point scale from 1 (limited severity perception) to 7 (greatest severity perception). For parent form item 4 (including 1.4, 2.4, and 4) where responses are given from greatest severity perception to most limited severity perception, the scale should be reversed to range from most limited to most severe for scoring purposes. That is, where the question is positively worded as "Effectively treating him/herself…" a response of "Extremely unlikely" should be scored as a 7 rather than a 1.

Domains for the FAIM are to be defined as follows:

Teen and adult form –

a) Expectation of outcomes (includes items 1-4)
b) Risk of Accidental Exposure (includes item 5)
c) Emotional Impact (includes item 6)

Child form –

a) Expectation of outcomes (includes items 1-4)
b) Risk of Accidental Exposure (includes item 5)
c) Emotional Impact (includes item 6)

Parent form teen –

a) Parent expectation of outcomes (includes all items on form 1-4)

Parent form 4-12 –

a) Parent expectation of outcomes (includes items 1.1-1.4)
b) Child expectations of outcomes (includes items 2.1-2.4)

Domain scores are the arithmetic average of the non-missing items comprising a domain. For child and teen/adult forms, a total score is defined as the arithmetic average of all the non-missing items. For any respondent and form, the domain and total scores should be calculated only if more than half the items comprising that domain/total score have valid responses.

For subjects 4-17, descriptive statistics for domain scores and for non-parent forms the total scores along with their changes from baseline will be tabulated. Data will be summarized for the safety population separately by questionnaire age group and responder (subject or parent).

Listing of the transformed scores, domain scores, and total scores will be provided, sorted by age group, group/cohort, and subject ID.

9.1.8. Treatment Satisfaction

Assessment of treatment satisfaction will be performed using the Treatment Satisfaction Questionnaire for Medication (TSQM-9) and an exit survey including palatability questions. Both are to be completed after the Exit DBPCFC.

The TSQM-9 is a widely used instrument to assess treatment satisfaction with medication in studies where patient reported side effects have a potential to interfere with the objectives of the study. The instrument consists of 9 questions that comprise 3 scales.

Responses to the 9 individual items will be presented using descriptive statistics. The scale scores (effectiveness, convenience, and global satisfaction) will be calculated and summarized using descriptive statistics for the safety population age group 4-17.

The Effectiveness scale includes items 1-3, the Convenience scale includes items 4-6, and the Global Satisfaction scale includes items 7-9. Each scale will be scored as: 100*[(sum of non-missing responses) minus the number of non-missing responses] divided by the maximum possible score of the sum of non-missing responses. If more than one item within the scale has a missing result then the scale score will not be calculated.

In addition to the TSQM-9, an exit survey will be performed at study exit. The survey includes questions on study drug palatability, frequency of taking study drug as instructed,

impact on attending clinic visits, interest in continuing to take study drug, if the subject would recommend the study drug to others, and burden of treatment. Responses to each item will be summarized for the safety population for each type of instrument administered (parent or subject ages 12 and older, and subjects ages 4 to 11 for the drug palatability as appropriate).

9.2. INTERIM ANALYSIS

There is no interim analysis of efficacy planned for this study.

10. SAFETY

Safety will be assessed based on extent of exposure, concomitant medications, physical examinations, and all the safety endpoints defined in Section 3.2.

Unless otherwise noted, safety data will be summarized descriptively and the Safety population will be used for all summaries of safety parameters. Event data such as adverse events will be summarized by study period in 2 parts. Part 1 includes Group 1 study periods of IDE and updosing, maintenance, and overall and Group 2 Cohort 2 periods of QOD, BIW, and overall. Part 2 includes all Group 2 QD AR101 rollovers and summarizes by Cohort 1, Cohort 3A, combined Cohorts 3B and 3C (and their respective periods of QD, QOD, BIW, and Overall), and Overall QD representing all Group 2 QD. In general, safety data will be summarized for age group 4-17 years. Additionally, key safety tables will be presented for 18-55 year old subjects. Safety listings will include all treated subjects, sorted by age group and group/cohort.

10.1. STUDY TREATMENT EXPOSURE

Study treatment exposure will be summarized by age group (ages 4 to 17 and 18 to 55), group/cohort, and study period. The calculation of exposure will be based on in clinic dosing data and the dose level of the dispensed study medication.

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

Study Period	Group/ Cohort	First Dose Date	Last Dose Date
Initial Dose Escalation	Group 1	Date of first in-clinic dose at Initial Dose Escalation Day 1	Date of last in-clinic dose at Initial Dose Escalation Day 1 or Day 2
Up-Dosing	Group 1	The day following the date study product was dispensed at Initial Dose Escalation Day 2	Date of first in-clinic dose of 300 mg at the End Up-Dosing 300 mg Visit. For subjects who do not reach the 300 mg dose: the latest of the last in-clinic dose during the Up-Dosing period or the date of last study drug administration as entered on the Early Termination CRF page.
Maintenance	Group 1	The day following the date study product was dispensed at End Up-Dosing 300 mg Visit	Latest of the last in-clinic dose during the Maintenance period or the date of last study drug administration as entered on the Early Termination CRF page
QD	Group 2, Cohorts 1, 3A, 3B, 3C	The first day of AR101 therapy in study ARC004 while on QD dosing	The last day of AR101 therapy in study ARC004 while on QD dosing
QOD	Group 2, Cohorts 2, 3B, and 3C	The first day of AR101 therapy in study ARC004 while on QOD dosing	The last day of AR101 therapy in study ARC004 while on QOD dosing

BIW	Group 2, Cohorts 2, 3B, and 3C	The first day of AR101 therapy in study ARC004 while on BIW dosing	The last day of AR101 therapy in study ARC004 while on BIW dosing

The total amount of study product consumed will be calculated as the sum of in-clinic doses plus the sum of doses taken at home as recorded in the 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 last dose of study product minus the date of the first dose of study product plus one during the study period. Duration of exposure will be summarized using descriptive statistics as well as categorically by the expected lengths of the overall treatment period.

- 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.

- For Group 1 only: Maximum dose achieved (mg/day) during up-dosing using descriptive statistics for continuous data as well as 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.

- For Group 1 only: 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 non-missing valid diary entries will be used to estimate at-home dosing compliance. The following measures of compliance with at-home dosing will be calculated:

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

At-home dosing data will be listed. Daily diary records, including date and time, whether a full or partial dose was consumed (or the dose was missed), reason for partial or missed dose will be listed.

10.2. PRIOR, CONCOMITANT, AND RESCUE MEDICATIONS AND THERAPIES

All medications recorded on the Concomitant Medications CRF 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 1 and Preferred Name. The highest ATC drug level available in the data will be presented, starting with ATC level 4 drug class. Level 3 will be presented if level 4 is missing, and level 2 will be presented if both level 3 and level 4 are missing. All prior and concomitant medications, all rescue medications, and epinephrine medications will be listed separately.

Prior medications are defined as those which are only taken prior to the date of the first dose of study product on Day 1 (i.e., medication end date is prior to the date of first dose of study product). For Group 2, prior medications are those prior medications entered on the concomitant medications page of ARC003. For Group 1, prior medications are those prior medications entered on the concomitant medications page of ARC004.

Concomitant medications are medications taken at any time during the ARC004 treatment period. Any medications recorded for which dosing began after the last dose of study treatment will also be classified as concomitant medications. As needed (PRN) medications, which may or may not be taken for long periods of time, but which are prescribed to the subject for a period that overlaps with 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. Unless administered during a DBPCFC, each use of a rescue medication during OIT should be associated with a corresponding adverse event (AE).

Prior medications and concomitant medications excluding rescue medications will be summarized by ATC Class, Preferred Name and group/cohort. Subjects will be counted no more than one time per Preferred Name and no more than one time per ATC Level 4 in the summary.

Rescue medications will be summarized by ATC Class and Preferred Name for the following group/cohort periods:

- Group 1: IDE and up-dosing, Maintenance, and Overall

- Group 2 Cohort 2: QOD, BIW, and Overall
- Group 2 Cohort 1 (includes QD only)
- Group 2 Cohort 3A (includes QD only)
- Group 2 Combined Cohorts 3 B and C: QD, QOD, BIW, and Overall
- Group 2 Overall QD
- Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, and Group 2 Cohort 3C by Maintenance and Exit DBPCFC

Concomitant non-drug therapies will be listed by subject.

10.3. ADVERSE EVENTS

Treatment-emergent adverse events (TEAEs) will be summarized excluding symptoms recorded during the food challenges. The AE summaries will report AEs and in-clinic dosing symptoms.

If symptoms are recorded as part of an anaphylactic reaction, only the single anaphylactic reaction event will be summarized and not the individual symptoms.

All reported adverse events (AEs) will be classified into System Organ Class (SOC) and Preferred Term (PT) using Medical Dictionary for Regulatory Activities (MedDRA) version 19.1. The 5-point CoFAR severity grading scale was used for coding allergic reactions. The 3-point Muraro grading scale was used for grading the severity of anaphylactic reactions. The 5-point CTCAE severity grading scale was used for coding all other adverse events.

In Group 1, treatment-emergent adverse events are defined as those AEs with onset after the first dose of study product in ARC004 and no more than 30 days after the last dose of study product. In Group 2, as all patients are assumed to be continuing AR101 from study ARC003, all reported adverse events with onset no more than 30 days after the last dose of study product will be defined as treatment-emergent. Non treatment-emergent AEs will be included in subject listings, but not summarized. An additional summary of age group 18-55 will be included for the safety population for the overall AE summary and SAE summary.

Summaries that are displayed by system organ class and preferred terms will be ordered by descending incidence of system organ class and preferred term within each system organ class. 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 TEAEs and treatment-emergent serious adverse events (TESAEs), subject incidence of TEAEs and TESAEs meeting various criteria and exposure-adjusted incidence rates of TEAEs and TESAEs 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 during the study period;

- Subject incidence of TEAEs by MedDRA system organ class and preferred term;

- Exposure-adjusted event rates for the most frequent TEAEs (i.e., TEAEs occurring in ≥ 5% of the Safety Population) by MedDRA preferred term;

- Subject incidence of treatment-related AEs by MedDRA system organ class, and preferred term;

- Exposure-adjusted event rates for the most frequent treatment-related AEs (i.e., treatment-related AEs occurring in ≥ 5% of the Safety Population) by MedDRA preferred term;

- Subject incidence of grade ≥ 3 severity TEAEs by MedDRA system organ class and preferred term;

- Subject incidence of grade ≥ 3 severity treatment-related AEs by MedDRA system organ class and preferred term;

- Subject incidence of TESAEs by MedDRA system organ class and preferred term;

- Subject incidence of TEAEs leading to discontinuation of study product by MedDRA system organ class and preferred term;

- Subject incidence of TEAEs associated with Epinephrine use by MedDRA system organ class and preferred term.

- Subject incidence of hypersensitivity TEAEs by MedDRA system organ class and preferred term, where hypersensitivity TEAEs are all TEAEs from the in-clinic dosing CRFs and all TEAEs indicated as an allergic reaction on the AE CRF

- Subject incidence of hypersensitivity treatment-related AEs by MedDRA system organ class and preferred term;

- Exposure-adjusted event rates for the most frequent hypersensitivity TEAEs (i.e., hypersensitivity TEAEs occurring in ≥ 5% of the Safety Population) by MedDRA preferred term;

- Subject incidence of TEAEs associated with nonstudy product food allergen exposure by MedDRA system organ class and preferred term; and

- Subject incidence of TEAEs leading to early withdrawal by MedDRA system organ class and preferred term

At each level of summarization (e.g., any AE, system organ class, and preferred term), subjects experiencing more than one TEAE will be counted only once within each study period. In the summary of TEAEs 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 (4-17 years and 18-55 years), group/cohort, subject, study period, and event. Serious TEAEs; Serious DBPCFC AEs; severe, life-threatening, or fatal adverse events; anaphylactic reactions; and AEs leading to discontinuation, reduction, or interruption of the study product will be presented in separate data listings.

10.4. FOOD ALLERGEN EXPOSURE

The occurrence of a safety event associated with accidental food allergen ingestion will be reported as a food allergen exposure. Any such event that meets the definition of an SAE will also be reported as an adverse event. All reported food allergen exposures will be listed by age group, group/cohort, and subject.

The number of subjects experiencing any food allergen exposure, the number of subjects experiencing a food allergen exposure in response to peanut (or nonpeanut), the number of exposures of each (peanut-related and nonpeanut related) experienced per subject, and the total number of food allergen exposures (peanut and nonpeanut related) will be summarized. The number of exposures considered SAEs, those that required treatment, and those that required epinephrine use will also be summarized.

10.5. SYMPTOMS DURING DOSING

During study product dosing, the severity of pre-specified allergy symptoms is rated as mild, moderate, severe, life-threatening, or death according to the CoFAR severity grading scale. In addition, the presence of dose-related 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 for each cohort. For Group 1, this will include all dose levels included in initial dose escalation, up-dosing, and 300 mg QD maintenance and extended maintenance dose. For Group 2, subjects are expected to dose at the 300 mg level, but dose reductions may occur. Any doses other than 300 mg for Group 2 will be summarized, but this data is expected to be sparse. Dosing at the study site and at-home dosing will be summarized separately.

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.

10.6. SYMPTOMS DURING DBPCFC

During each food challenge, the severity of pre-specified symptoms is rated on the CoFAR scale as mild, moderate, severe, life-threatening, or death at each dose level of each food product. In addition, the presence of any dose related symptoms is identified.

The number of subjects experiencing any dose related symptoms and the maximum severity of any symptoms will be summarized by individual dose level and overall during the Screening peanut challenge, Maintenance peanut challenge (Group 1 only) and the Exit peanut challenge by group/cohort. Subjects will be counted at most once per type of challenge and dose level for symptom severity at the most severe level recorded for that subject.

Symptoms at the maintenance and Exit DBPCFCs will be listed by age group, group/cohort, and subject.

10.7. PREGNANCY TEST RESULTS

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

10.8. SPIROMETRY AND PEF

Spirometry and/or Peak Expiratory Flow (PEF) assessments are performed at screening, initial dose escalation, and prior to any DBPCFC. PEF is also performed at each Up-dosing, maintenance, and extended maintenance visit. Spirometry may be performed at any time during the study where a subject's pulmonary status is in question. Three attempts of PEF are performed and the best (highest) value flagged in data listings. Only the best PEF value will be summarized.

Observed values for PEF, forced expiratory volume in one second (FEV_1), FEV_1 percent predicted, FEV_1/FVC ratio where FVC is forced vital capacity, and FEV_1/FVC percent predicted as well as changes from baseline will be summarized at each applicable visit and by group/cohort. Group 1 will be summarized separately for up-dosing and maintenance. The extended Maintenance summary will include Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, and Group 2 Cohort 3C.

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

10.9. VITAL SIGNS

BMI will be calculated as (weight in kilograms) / (height in meters)^2.

Vital signs (pulse rate, systolic/diastolic blood pressure, body temperature, height, BMI and weight) will be listed by age group, group/cohort, subject and visit. Observed values and change from baseline will be summarized by group/cohort at each scheduled visit and

time point for pulse rate, systolic/diastolic blood pressure, and body temperature. At DBPCFC vital signs are scheduled to be taken just prior to each dose given. At Initial Dose Escalation, Up-dosing, and Maintenance visits, and EM visits, vital signs are to be taken pre-dose and after each dose given. If multiple measurements are recorded pre-dose or post-dose at a given in-clinic visit, the last pre-dose measurement and the first post-dose measurement will be used in the summaries.

10.10. PHYSICAL EXAMINATION

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

Missing data will not be imputed.

10.11. ASSESSMENT OF ASTHMA CONTROL

Assessment of asthma control in asthmatic subjects using the Asthma Control Test (ACT) questionnaire will be performed.

For subjects 12 years old 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 12, there are 4 questions for the subject and 3 questions for the parents to complete. Subject responses range from 0 (least control) to 3 (greatest control). Parent responses range from 0 (every day) to 5 (no days). The sum of all 7 questions will make up the total score. The total ACT score for subjects under 12 will range from 0 (least control) to 27 (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 questions have a missing response, the total ACT score will not be calculated for that subject.

All analyses of the ACT will be performed separately by subject age group (4-11 for subject and parent, 12-17 subjects). Summary statistics of the score for question, total score and change from baseline will be tabulated by visit and group/cohort. A shift table of asthma control (adequate, not adequate, missing) will be summarized by group/cohort 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, group/cohort, subject, and visit.

10.12. ASSESSMENT OF GI SYMPTOMS BY PEESS

Subjects who discontinue treatment due wholly or in part to GI AEs will be instructed to complete the Pediatric Eosinophilic Esophagitis Symptom Scores (PEESS v2.0)

questionnaire (Martin, 2015; Franciosi, et al., 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.

PEESS results including the frequency total, severity total, and total scores, will be listed.

10.13. EPINEPHRINE USE AS RESCUE MEDICATION

Epinephrine use is defined as any rescue medication with a preferred name of 'EPINEPHRINE' when coded as described in Section 10.2.

All subjects, per protocol, are required to have epinephrine autoinjectors for use in case of a suspected anaphylactic reaction occurring outside of the clinic. There are, however, differences in how, and even if, physicians record the prescription of epinephrine autoinjectors for as-needed (PRN) use. As a consequence of this, the presence or absence of a PRN prescription for epinephrine cannot be taken to indicate epinephrine usage, regardless of whether the prescription was written prior to, or after, enrollment in the study. What is important is to be able to quantitate 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 on the Rescue Medication CRF form. In cases where a PRN epinephrine prescription is issued after the start of study-product dosing, the sites will be queried as to if, and when, epinephrine was actually administered and to treat what specific event.

All epinephrine use will be listed for the Safety population.

In addition, epinephrine use excluding food challenges will be summarized by group/cohort and study period. The total number and percent of subjects with any episode, 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). For these subjects, demographics including age, age category, baseline PS IgE, and baseline PS IgE category will be included. 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.

10.14. ANAPHYLACTIC REACTIONS

All reported anaphylactic reaction events will be listed by age group, group/cohort, and subject.

Each anaphylactic reaction will be identified by the following triggers:

- DBPCFC
- Study product
- Food allergen other than study product
- Other

Anaphylactic reactions will be summarized in the Safety population separately for the following group/cohort study periods:

- Group 1: IDE and Up-Dosing, Maintenance, Overall
- Group 2 Cohort 2: QOD, BIW, and Overall
- Group 2 Cohort 1 (includes QD only)
- Group 2 Cohort 3A (includes QD only)
- Group 2 combined cohorts B and C: QD, QOD, BIW, and Overall
- Group 2 Overall QD
- Group 1, Group 2 Cohort 1, Group 2 Cohort 2, Group 2 Cohort 3A, Group 2 Cohort 3B, and Group 2 Cohort 3C by Maintenance and Exit DBPCFC

By each group/cohort and study period, the summary will include the number of anaphylactic reactions; the number of anaphylactic reactions by trigger (study product, food allergen other than study product, and other); and subject incidence of the total number of subjects experiencing an anaphylactic reaction, subjects experiencing an anaphylactic reaction by number of episodes, subjects experiencing an anaphylactic reaction by maximum severity using the Muraro grading scale (Muraro, 2007; Muraro, 2014), subjects experiencing an anaphylactic reaction that was a serious adverse event, subjects experiencing an anaphylactic reaction that required epinephrine use, subjects experiencing an anaphylactic reaction by location (home or study site), and the individual symptoms involved.

If any symptom associated with an anaphylactic reaction is serious, the reaction will be classified as serious. This is in addition to, not instead of, the Muraro grade. If any symptom associated with an anaphylactic reaction is related to study product, then the reaction will be classified as related to study product.

11. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

The analysis methods and summaries detailed in the SAP include, but are not limited to, all analyses included in the Statistical Considerations section of the protocol, with the following exceptions:

- Baseline for DBPCFC is defined as the Exit food challenge from ARC003 for both Group 1 and Group 2 subjects.

- For secondary endpoints relating to post-maintenance DBPCFC, only Group 1 post-maintenance DBPCFC results will be summarized. Group 2 Exit ARC003 food challenge data will not be summarized in ARC004 as post-maintenance.

- Post-maintenance DBPCFC as referenced in the protocol will be referred to as maintenance DBPCFC in this document and all corresponding outputs.

Any deviations from the plans detailed in this SAP will be described and justified in the final clinical study report. A separate document to this SAP will provide a table of contents and mockups for the expected layout and titles of the TLFs. Any changes to format, layout, titles, numbering, or any other minor deviation will not necessitate a revision to the SAP nor will it be considered a deviation from planned analyses. Only true differences in the analysis methods or data handling will necessitate such documentation.

12. REFERENCE LIST

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.

13. PROGRAMMING CONSIDERATIONS

All tables, data listings and figures (TLFs) will be generated using SAS® for Windows, Release 9.4 or higher (SAS® Institute Inc., Cary, NC, USA).

Report summaries will be generated using validated Base SAS® software, version 9.4 or higher, on a PC or server-based platform. Additional validated software may be used to generate analyses, as needed.

Data will be analyzed by Precision biostatistics personnel. Statistical analyses will be reported with tables, figures and listings, presented in rich text format and using recommended ICH numbering. Output specifications for all tables, figures and listings will be in conformance with guidelines specified by the ICH in Appendix 7 of the *Electronic Common Technical Document Specification* (Apr 2003).

14. QUALITY CONTROL

All SAS programs that create outputs or supporting analysis datasets will be validated by a second statistical programmer or biostatistician. At a minimum, validation of programs will consist of a review of the program log, review of output or dataset format and structure, and independent confirmatory programming to verify output results or dataset content. Additionally, all outputs will undergo a review by a senior level team member before finalization.

The content of the source data will be reviewed on an ongoing basis by project statistical programmers and statisticians. Data will be checked for missing values, invalid records, and extreme outliers through defensive programming applications, analysis-based edit checks, and other programmatic testing procedures. All findings will be forwarded to the project data manager for appropriate action and resolution.

15. STUDY SCHEDULE

Refer to the protocol for the full study schedule of events (Appendix 1).

16. INDEX OF TABLES, LISTINGS AND FIGURES

An index of the planned statistical outputs will be provided in the shell TLF document.