Challenging to Food With Escalating Thresholds for Reducing Food Allergy
Scott H Sicherer, M.D
NCT03907397
Document Date: 1/19/2024

Statistical Analysis Plan

ASTHMA AND ALLERGIC DISEASES COOPERATIVE RESEARCH CENTERS

Immune Basis & Clinical Implications of Threshold-Based Phenotypes of Peanut Allergy

ChAllenging to Food with Escalating ThrEsholds for ReducIng Food Allergy (CAFETERIA)

Sponsored by the National Institute of Allergy and Infectious Diseases (NIAID)

Version 1.0
Version Date: January 2024

MODIFCATION HISTORY

Version Number	Date of Document Version	Significant Changes from Previous Authorized Version
1.0	1/23/2024	Not applicable – first version

TABLE OF CONTENTS

ABBREVIATIONS AND DEFINITIONS

DBPCFC	Double-blind, placebo-controlled oral food challenge
DSMB	Data Safety Monitoring Board
MAR	Missing at random
MNAR	Missing not at random
OFC	Oral food challenge
SU	Sustained unresponsiveness

PURPOSE OF STATISTICAL ANALYTICAL PLAN (SAP)

The purpose of this SAP is to outline the planned analyses to be completed for the "ChAllenging to Food with Escalating ThrEsholds for Reducing Food Allergy" (CAFETERIA) trial (protocol number AADCRC-ISMMS-03). The SAP is based on protocol [clinicaltrials.gov ID NCT03907397] revision 3.0 finalized November 2021. This SAP covers study endpoints related to the primary and secondary clinical objectives and does not address the primary and secondary mechanistic objectives. The analyses identified in this SAP will be included in abstracts and manuscripts reporting the results of the trial. Exploratory analyses not necessarily identified in this SAP may also be performed. Any post hoc, or unplanned, analyses not explicitly identified in this SAP will be clearly identified as such in any published papers from this study. This SAP may be updated in response to additional developments, either within or outside the trial. All revisions will be made prior to the data lock and the primary analysis.

1. INTRODUCTION

Food allergy affects up to 5% of adults and 4-8% of young children in the US, an estimated 15 million Americans.[1,2] Food allergy carries a significant cost, and a negative impact on nutrition and quality of life.[3-6] Peanut allergy affects nearly 2% of children, is often severe, and life-long.[7] The current standard for food allergy management relies on strict avoidance of the offending food. Children are monitored for resolution of food allergy with periodic oral food challenges (OFC), which are performed when diagnostic tests and the medical history suggest a possibility that the allergy has resolved. An OFC involves gradually feeding increasing amounts of the food under medical supervision until a "meal-sized" portion is ingested, unless there are symptoms requiring cessation of dosing. These tests are performed by allergists for clinical diagnosis when they suspect resolution of peanut allergy or sensitization without reactivity, using office-measured foods such as peanut butter, peanut flour, or Bamba, in amounts measured by common kitchen materials such as spoons.[8] A child who reacts at any dose during an OFC, even at the top dose, even with mild symptoms, is considered allergic and advised to continue strict avoidance until the next OFC is performed, typically not less than a year later.[4,9-11]

However, the paradigm of strict avoidance of a food as therapy for food allergy is changing. We believe that the next logical step is a simple and cost-effective approach of allowing children with high threshold peanut allergy to ingest tolerated amounts of peanut. The approach is based in part on our prior success with allowing children to ingest milk and egg in tolerated forms, and the promise seen in OIT studies[12-20].

Currently, allergists perform OFCs in the office setting by feeding supermarket forms of peanut measured with simple kitchen materials often using foods brought in by the patient. We propose that allowing children with high threshold peanut allergy to ingest tolerated, sub-threshold amounts of peanut, using home-purchased, home-measured foods based on the results of an OFC, may be associated with benefits such as further increased threshold with time and potentially sustained unresponsiveness (SU, remission off daily ingestion), should be safe, and result in improved quality of life. This approach may become a prototype for studying additional foods. Additionally, we will undertake immunologic genomic, and transcriptomic characterization of a high threshold endotype of peanut allergy to inform identification of biomarkers and mechanisms of threshold, response to therapy, reaction severity, and SU/tolerance/remission.

Therefore, we have designed a trial to determine if there is a benefit to allowing children with high threshold peanut allergy to ingest home measured amounts of peanut below their threshold, as compared to standard care avoidance. This document serves as the SAP for the CAFETERIA trial.

2. STUDY OBJECTIVES

2.1 Primary Clinical Objective

The primary objective of this study is to determine whether allowing ingestion of sub-threshold amounts of peanut in those with a high threshold (tolerate at ≥143 mg peanut protein but <5043 mg peanut protein on supervised double-blind, placebo-controlled oral

food challenge [DBPCFC]) will be associated with attaining even higher thresholds over time compared to those avoiding peanut.

2.2 Secondary Clinical Objectives

The secondary clinical objectives include assessing the development of sustained unresponsiveness (a surrogate term for tolerance without daily ingestion), effects on quality of life, changes in skin prick test mean wheal diameter and safety in participants randomized to peanut ingestion compared to avoidance.

3. STUDY OVERVIEW

3.1 Study Design

This is a prospective two-arm, parallel-group, randomized (1:1) controlled open trial of a diet allowing ingestion of tolerated, home-purchased, home-measurable quantities of peanut in children allergic to peanut in higher amounts.

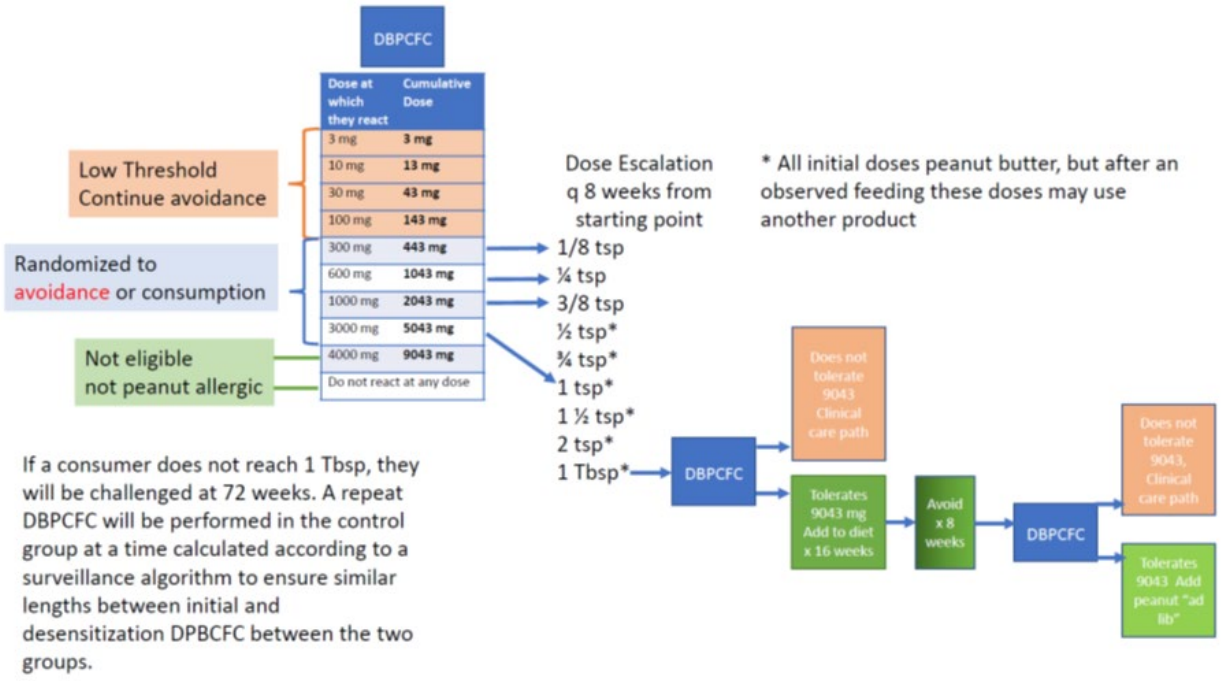
3.1.1 Study Duration and Time Points

Participants in the intervention group may begin with different starting amounts of peanut butter, depending upon the baseline reaction threshold (Figure 1). Every 8 weeks, participants will return for attempting to ingest a higher amount of peanut along a ladder of 1/8 tsp, ¼ tsp, 3/8 tsp, ½ tsp, ¾ tsp, 1 tsp, 1.5 tsp, 2 tsp, 3 tsp. A repeat DBPCFC (desensitization DBPCFC visit; primary endpoint visit) will be performed 8 weeks after reaching 1 tablespoon (or equivalent peanut product after exceeding 3/8 teaspoon; see Protocol Section 3.1.3), or at 72 weeks. Subjects tolerating the full challenge amount will add peanut to the diet for 16 weeks and then avoid peanut for 8 weeks, followed by a DBPCFC to assess for sustained unresponsiveness.

Participants in the control arm will receive monthly telephone follow up. They will have a visit 16 weeks following the baseline DBPCFC for review of avoidance instructions and collection of a blood sample, saliva and stool for mechanistic studies, and quality of life and diet questionnaires. They will have a repeat DBPCFC (desensitization DBPCFC visit; primary endpoint visit) to peanut at a time determined by a surveillance algorithm (see Protocol Section 13.3) to ensure similar lengths of time between initial and repeat DBPCFC between the two randomization groups. In brief, after the baseline DBPCFC, patients in the avoider group will be assigned a 'time to the second DBPCFC' as a value randomly selected from the distribution of the consumer patients with the same baseline reactive dose. Initial distributions for time to DBPCFC will be assumed to follow a uniform distribution centered on the treatment expectations anticipated by the clinicians based on initial DBPCFC (i.e. ~32 weeks for those who fail at 3000 mg, ~72 weeks for those who fail at 300 mg) plus minus 3 weeks. After the recruitment of every 10 patients, the time to OFC distribution for each baseline reactive dose will be updated as a uniform distribution derived from the time to OFC of consumer patients if 2 or more

observations are available, otherwise we will keep the initial uniform distribution for that baseline reactive dose.

Figure 1: Study Scheme



3.1.2 Randomization and Masking

Children age 4-14 years with high threshold peanut allergy will be randomly assigned (1:1) to ingest a sub-threshold amount of peanut daily or to follow avoidance. Randomization will be stratified by age (4 - <10 years of age vs. 10 – 14 years of age) and cumulative reaction dose at baseline DBPCFC (443mg; 1043 mg; 2043 mg; 5043 mg). The data and statistical coordinating center (DSCC) will maintain control of stratification and randomization. Randomization will be performed centrally through a Web-based data collection system that automates the delivery of the randomization assignments. Neither the participant nor the investigators will be blinded to randomization assignment. The study staff conducting the DBPCFCs, however, will be blinded to the participant randomization status. Trial oversight will be provided by an independent Data and Safety Monitoring Board (DSMB).

4. ANALYSIS POPULATIONS

Four populations will be used for all summaries and analyses.

Screened Population

The screened population will consist of all subjects who signed a consent form and met enrollment eligibility prior to the baseline DBPCFC.

Intent-to-Treat (ITT) Population
The ITT population will consist of all randomized subjects grouped by their assignment at randomization whether or not they actually received the treatment to which they were assigned. This sample will be used for summaries and analyses of the primary and secondary (where applicable) clinical endpoints. The ITT population will be used for the primary and secondary (where applicable) analyses and description of patient and baseline characteristics.

All Available Data Population
The All Available Data population will consist of randomized subjects with at least one longitudinal assessment of the secondary clinical endpoint (e.g., Quality of Life). Subjects will be analyzed in the group they were randomized. The All Available Data population will be identical to the ITT population if all randomized subjects have at least one measurement.

Per Protocol Population
The Per Protocol Population will consist of subjects in the intervention group who attended two-thirds of the updosing visits and the desensitization DBPCFC visit and subjects in the control group who did not consume peanut based on IgG 4 levels.

Safety Population
The Safety population will consist of all randomized subjects in the intervention group who received any amount of peanut after the baseline DBPCFC and all randomized subjects in the control group. The Safety population will be identical to the ITT population if all randomized subjects in the intervention group receive the assigned treatment. The Safety population will be used for the analysis of safety data.

5. STUDY ENDPOINTS

5.1 Primary Clinical Endpoint
The primary endpoint is success on the desensitization DBPCFC defined as the ability tolerate a dose at least 2 steps higher than the baseline DBPCFC or 9043 mg of peanut protein.

5.2 Secondary Clinical Endpoints
The following secondary clinical endpoints will be assessed:

1. Sustained unresponsiveness (intervention group) or natural tolerance (control group; attaining tolerance without an intervention)
2. Safety parameters (acute allergic reactions, including anaphylaxis, gastrointestinal side effects)
3. Quality of life measures
4. Skin prick test mean wheal size

5.3 Exploratory Clinical Endpoint
1. Severity Score for acute allergic reactions

6. STATISTICAL METHODOLOGY

6.1 General Principles

Continuous variables will be summarized using the following descriptive statistics: number of non-missing values, means, standard deviations, medians, interquartile range, maximum, and minimum. Categorical variables will be summarized using number of non-missing values, counts and percentages.

For any variable measured at multiple points in time, change from baseline will be calculated as the difference between the value of the variable at a specific point in time (e.g. 1 year) minus the baseline value. Relative change from baseline will be calculated as the value of a parameter at a specific point in time minus the baseline value of the parameter divided by the baseline value of the parameter. Percent change will be calculated as the relative change multiplied by 100.

All hypothesis testing will be conducted at the 0.05 two-sided significance level unless otherwise specified. P-values will be rounded to three decimal places. P-values less than 0.001 will be reported as <0.001 in tables. P-values greater than 0.999 will be reported as >0.999.

Should any of the statistical methods proposed prove unsuitable during data analysis, more appropriate methods will be used. These include data transformation (for example to a logarithmic scale) to satisfy model assumptions such as normally distributed residuals with constant variance, the application of non-parametric techniques or the use of a different link function or modeling technique.

Additional ad-hoc analyses may be conducted as deemed appropriate.

All analyses will be conducted using SAS V9.4 or higher and R statistical software (version 4.2.2 or higher).

6.2 Missing Data

6.2.1 Missing Baseline Data

Missing baseline values will not be imputed and summaries will be based on all available data.

6.2.2 Missing Primary Outcome Data

The handling of missing outcome values in the primary endpoint is described in section 6.5.1.2.

6.2.3 Missing Secondary Clinical Outcomes Data

The handling of missing outcome values in the secondary clinical endpoints is described in section 6.5.2.

6.3 Crossover

Crossovers are subjects who after randomization switch from the allocated treatment to the non-allocated treatment. Subjects randomized to peanut ingestion who never consume peanut after the baseline DBPCFC or who permanently discontinue peanut consumption will be considered crossovers. In addition, subjects randomized to avoidance who start ingesting peanut and continue consumption through the desensitization DBPCFC visit (or study discontinuation, whichever comes first) will be considered to have crossed over. Should crossover occur, crossovers will be analyzed as belonging to the group to which they were randomized according to the ITT principle in all efficacy analyses. Safety analyses will be conducted according to the definition of the Safety population described in section 4.

6.4 Patient Characteristics

6.4.1 Patient Disposition

Disposition will be summarized in the screened and ITT populations.

Disposition summaries of the screened population will include:

- The number of consented subjects eligible for the baseline DBPCFC
- The number and percentage of consented subjects who were not randomized and reasons (e.g., subject ingested <143 mg peanut protein; subject ingested ≥5043 mg peanut protein; subject did not complete baseline food challenge visit)
- The number and percentage of consented subjects randomized

Disposition in the ITT population will be summarized by randomization group and will include:

- The number of subjects randomized
- The number and percentage of subjects who received their assigned procedure
- The number and percentage of subjects withdrawn or lost to follow-up by the desensitization DBPCFC visit (primary endpoint visit) and the primary reason for withdrawals
- The number and percentage of subjects withdrawn or lost to follow-up by the sustained unresponsiveness visit in the intervention group and the primary reason for withdrawals

6.4.2 Patient Characteristics

6.4.2.1 Demographic characteristics

Demographics including age, sex, race and ethnicity will be summarized by randomization assignment using the appropriate descriptive statistics.

6.4.2.2 Baseline characteristics

Baseline characteristics will be summarized by randomization assignment using the appropriate descriptive statistics. The specific baseline variables collected are detailed in the protocol and include baseline cumulative reaction dose (443 mg; 1043 mg; 2043 mg;

5043 mg), medical/allergy history, physical exam findings, vital signs, spirometry, skin prick test to peanut blood samples, quality of life survey, diet questionnaire, urine pregnancy if child bearing potential, and saliva and stool collection.

6.5 Primary and Secondary Clinical Endpoint Analyses

6.5.1 Analysis of the Primary Clinical Endpoint and Determination of Sample Size

6.5.1.1 Determination of Sample Size

For sample size calculations, we assume a conservative estimate for the proportion of spontaneous tolerance in the control group of 10%. We believe that an additional absolute increase of 45 percentage points under the intervention is feasible and clinically meaningful (55% vs. 10%). We also assume a drop-in rate (avoiders to consumers) and drop-out rate (consumers discontinuing study treatment) of 5%and 20%, respectively. This results in an attenuated effect size of 33.75% percentage points (46% vs. 12.25%. A total of 72 children randomized with equal probability to the active or control arm (36 per group) provides approximately 85% power to detect a difference of 33.75% (46% versus 12.25%) in the proportion of children who tolerate at least 2 steps higher amount from baseline or the full amount of peanut protein by the desensitization DBPCFC visit, based on a 0.05 level continuity corrected chi-squared test. We believe the drop-in rate will be kept at a minimum since parents and children will likely not attempt dose escalation on their own. Peanut-specific IgG levels will be monitored in all participants and those with rising peanut-specific IgG from ingestion could be identified. The drop-out rate is conservative, based on our prior experience, which reported a drop-out rate of 15%.

6.5.1.2 Analysis of Primary Clinical Endpoint

The primary clinical endpoint of this study is the ability to tolerate a dose at least 2 steps higher than the one tolerated at baseline or 9043 mg of peanut protein by the desensitization DBPCFC. The null hypothesis is that there is no difference in the probability of participants who can consume higher amounts of peanut between subjects randomized to ingestion compared to subjects randomized to avoidance. The proportion of subjects in the intervention and the control group meeting the primary outcome will be compared using a continuity corrected chi-squared test at the 0.05 significance level.

As the intent-to-treat principle includes all randomized subjects in the primary analysis, patients who are missing the trial's primary endpoint will have their outcome imputed using multiple imputation assuming that the data are missing at random (MAR). The imputation model will be stratified by randomization assignment and include age, sex, baseline reactive dose strata, clinical severity during baseline DBPCFC as well as peanut-specific IgE levels (log transformed) at baseline and week 16. Since the model includes a mixture of variables types (i.e. continuous and discrete), a fully conditional specification method will be used.[21] The imputation process will be repeated 30 times to achieve maximal stability of

the procedure. A separate analysis will be conducted for each completed-and-imputed dataset. Rubin's rule[22] will be used to combine the 30 analyses and test the difference in success rates between the intervention and control groups.

6.5.1.3 Sensitivity Analysis of the Primary Clinical Endpoint

Sensitivity analyses will be performed to assess the sensitivity of the primary analysis results to choice of statistical test and missing imputation procedure as well as departures of the MAR assumption and to treatment cross-overs, if necessary. Sensitivity analysis include:

a. Stratified Chi-square test (by baseline reactive dose and age strata)
b. Fisher's exact test under 5 ad-hoc imputation strategies: i) all are failures, ii) are all failures except the highest dose at baseline which are successes, iii) all are failures except highest two doses at baseline, iv) all are successes except lowest dose at baseline, and v) all are successes. That is, the first and last of the five are the least conservative, respectively, and the others are in between but use the baseline dose as a means of an intermediate approach, which also incorporates a possible difference in distribution of baseline doses between the groups.
c. If the MAR assumption is not tenable, statistical methods that assume missing not at random (MNAR) such as pattern mixture models for non-ignorable missing data will be considered.
d. When cross-over rates are high, the ITT effect of being assigned to treatment may differ from the effect of actually receiving treatment. If cross-over rates are high, instrumental variable analysis will be performed to obtain an unbiased estimate of the treatment effect, using randomization assignment as the instrument.
e. Given the potential impact of COVID-19 pandemic on subject's return for the desensitization DBPCFC visit, a logistic regression model examining the effect of peanut ingestion on success of the primary endpoint will be conducted and adjusted for time from randomization to desensitization DBPCFC visit and whether patients were randomized prior to study pause in March 2020.

6.5.1.4 Examination of Subgroups of the Primary Clinical Endpoint

There are no subgroup analyses planned for the primary clinical endpoint.

6.5.2 Analyses of Secondary Clinical Endpoints

6.5.2.1 Sustained unresponsiveness or natural tolerance

The proportion of subjects who experience sustained unresponsiveness in the intervention arm (ability to consume 9043 mg of peanut protein at the SU DBPCFC visit) will be compared to the proportion of subjects who achieve natural tolerance in the control arm (ability to consume 9043 mg of peanut protein) in an intent-to-treat analysis using a continuity corrected chi-squared test. Given success on the desensitization DBPCFC, we anticipate that missing data for the SU DBPCFC visit

will be extremely low. However should the unexpected occur, missing values will be imputed as failures.

6.5.2.2 Immunological parameters

The skin prick test mean wheal diameter will be defined as the difference between the mean wheal diameter for the undiluted peanut extract and the negative control. Changes in skin prick test mean wheal diameter will be compared between randomization groups using linear mixed effects models for data collected at baseline and the desensitization DBPCFC visit. The analysis will include a random subject effect and the fixed effect of randomization assignment, continuous time, and their interaction. Spline models may be considered if the curves exhibit non-linear trends.

Missing data of the skin prick endpoint will not be imputed. The linear mixed effects model has the advantage that the estimation of the model parameters will be unbiased even in the presence of missing outcomes, assuming that the missing values depend only on the observed values (MAR). If the MAR assumption is not plausible, pattern-mixture modeling (which stratifies subjects by their pattern of missing data) will be used. The All Available Data population will be used for this analysis.

6.5.2.3 Quality of life

Changes in quality of life over the study as measured by The Food Allergy Quality of Life-Parental Burden[23] will be compared between randomization groups and analyzed using linear mixed effects models for data collected at baseline, week 16, and the desensitization DBPCFC visit. The analysis will include a random subject effect and the fixed effect of randomization assignment, continuous time, and their interaction. Spline models may be considered if the curves exhibit non-linear trends. If the MAR assumption for linear mixed models is not plausible, pattern-mixture modeling will be used. The All Available Data population will be used for this analysis.

6.5.2.4 Adverse Events

Individual adverse events will be summarized as the number (%) of events and number and (%) of patients with the event in peanut ingestion and avoidance subjects. Adverse events will be modelled using Poisson regression and the rate of individual adverse events will compared between randomization arms through the desensitization DBPCFC visit. Long-term adverse events will be evaluated for those in the intervention group who are tested for sustained unresponsiveness following the desensitization DBPCFC visit. The Safety population will be used for this analysis.

6.5.3 Analyses of Clinical Exploratory Endpoint

6.5.3.1 Severity Score for acute allergic reactions

Severity scores during OFCs will be calculated based on the severity grading system for allergic reactions developed by Dribin and colleagues[24] and will be compared

between randomization groups using generalized mixed effects models.

6.6 Multiplicity Adjustment

There will be no formal correction of the Type I error rate for multiple testing of statistical hypotheses for any of the secondary clinical endpoints. Therefore, p-values generated from these secondary analyses should be treated as descriptive in nature.

6.7 Data Lock

The dataset for the primary outcome analysis will be locked when all data through the last desensitization DBPCFC visit have been entered, all queries have been resolved, and data management processes have been completed. The entire database will be locked when all data for SU DBPCFC visit have been entered and all queries have been addressed.

REFERENCES

1. Sicherer SH, Sampson HA. Food allergy: recent advances in pathophysiology and treatment. AnnuRevMed 2009;60:261-77.
2. Gupta RS, Springston EE, Warrier MR, et al. The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics 2011;128:e9-17.
3. Lieberman JA, Sicherer SH. Quality of life in food allergy. Curr Opin Allergy Clin Immunol 2011;11:236-42.
4. Sampson HA, Aceves S, Bock SA, et al. Food allergy: a practice parameter update-2014. J Allergy Clin Immunol 2014;134:1016-25 e43.
5. Mehta H, Ramesh M, Feuille E, Groetch M, Wang J. Growth comparison in children with and without food allergies in 2 different demographic populations. J Pediatr 2014;165:842-8.
6. Nowak-Wegrzyn A, Groetch M. Nutritional aspects and diets in food allergy. Chem Immunol Allergy 2015;101:209-20.
7. Sicherer SH, Sampson HA. Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. J Allergy Clin Immunol 2018;141:41-58.
8. Bird JA, Groetch M, Allen KJ, et al. Conducting an Oral Food Challenge to Peanut in an Infant. J Allergy Clin Immunol Pract 2017;5:301-11 e1.
9. Boyce JA, Assa'ad A, Burks AW, et al. Guidelines for the Diagnosis and Management of Food Allergy in the United States: Summary of the NIAID-Sponsored Expert Panel Report. J Allergy Clin Immunol 2010;126:1105-18.
10. Muraro A, Werfel T, Hoffmann-Sommergruber K, et al. EAACI food allergy and anaphylaxis guidelines: diagnosis and management of food allergy. Allergy 2014;69:1008-25.
11. Nowak-Wegrzyn A, Assa'ad AH, Bahna SL, et al. Work Group report: oral food challenge testing. J Allergy Clin Immunol 2009;123:S365-83.
12. Leonard SA, Sampson HA, Sicherer SH, et al. Dietary baked egg accelerates resolution of egg allergy in children. J Allergy Clin Immunol 2012;130:473-80 e1.
13. Leonard SA, Caubet JC, Kim JS, Groetch M, Nowak-Wegrzyn A. Baked milk- and egg-containing diet in the management of milk and egg allergy. J Allergy Clin Immunol Pract 2015;3:13-23; quiz 4.
14. Nowak-Wegrzyn A, Bloom KA, Sicherer SH, et al. Tolerance to extensively heated milk in children with cow's milk allergy. The Journal of Allergy and Clinical Immunology 2008;122:342-7, 7.
15. Shreffler WG, Wanich N, Moloney M, Nowak-Wegrzyn A, Sampson HA. Association of allergen-specific regulatory T cells with the onset of clinical tolerance to milk protein. J Allergy Clin Immunol 2009;123:43-52 e7.
16. Wang J, Lin J, Bardina L, et al. Correlation of IgE/IgG4 milk epitopes and affinity of milk-specific IgE antibodies with different phenotypes of clinical milk allergy. J Allergy Clin Immunol 2010;125:695-702, e1- e6.
17. Kim JS, Nowak-Wegrzyn A, Sicherer SH, Noone S, Moshier EL, Sampson HA. Dietary baked milk accelerates the resolution of cow's milk allergy in children. J Allergy Clin Immunol 2011;128:125-31.
18. Caubet JC, Nowak-Wegrzyn A, Moshier E, Godbold J, Wang J, Sampson HA. Utility of casein-specific IgE levels in predicting reactivity to baked milk. The Journal of Allergy and Clinical Immunology 2012.
19. Ford LS, Bloom KA, Nowak-Wegrzyn AH, Shreffler WG, Masilamani M, Sampson HA. Basophil reactivity, wheal size, and immunoglobulin levels distinguish degrees of cow's milk tolerance. J Allergy Clin Immunol 2013;131:180-6 e1-3.
20. Caubet JC, Nowak-Wegrzyn A, Moshier E, Godbold J, Wang J, Sampson HA. Utility of casein-specific IgE levels in predicting reactivity to baked milk. The Journal of Allergy and Clinical Immunology 2013;131:222-4.

21. Berglund, P., Heeringa, S. Multiple imputation of missing data using SAS: SAS Institute, Cary, N.C.; 2014.
22. D.B. R. Multiple imputation for nonresponse in suveys: Wiley & Sons, Inc., New York; 1987.
23. Cohen BL, Noone S, Munoz-Furlong A, Sicherer SH. Development of a questionnaire to measure quality of life in families with a child with food allergy. J Allergy Clin Immunol 2004;114:1159-63.
23. AH, Risma KA, Castells M, Schneider LC, Wang J, Lee J, Mistry RD, Vyles D, Pistiner M, Witry JK, Zhang Y, Sampson HA. Severity grading system for acute allergic reactions:
A multidisciplinary Delphi study. J Allergy Clin Immunol 2021;148(1):173-181.

ADDENDUM

This is an addendum to the Statistical Analysis Plan for the CAFETERIA trial. The addendum outlines the statistical analyses that will be performed for the mechanistic studies conducted within the parent trial.

Mechanistic Objectives and Endpoints

The primary mechanistic objectives are to elucidate immune mechanisms induced by daily ingestion of sub-threshold amounts of peanut, and to identify biomarkers of and functional pathways underlying desensitization potential. We will test if peanut ingestion is associated with the following immune and transcriptomic processes:

- Suppression of serum peanut-specific IgE
- Increase in serum peanut-specific IgG and IgG4
- Change in peanut IgE and IgG4 epitope binding scores
- Decrease in basophil activation tests
- Decrease in peanut-specific multifunctional Th2 cells
- Increase in peanut-specific regulatory T cells
- Change in peripheral blood transcriptional signatures
- Up and downregulation of distinct gene ontology pathways

Primary Mechanistic Endpoints

1. Change in peanut and Ara h 2-specific IgE from baseline as measured by ImmunoCAP assay
2. Change in peanut and Ara h 2-specific IgG4 from baseline as measured by ImmunoCAP assay
3. Change in peanut epitope-specific IgE and IgG4 binding score from baseline
4. Change in basophil activation from baseline, as measured by %CD63+ basophil by flow cytometry/mass cytometry
5. Change in frequency of peanut-specific Th2 cells, as measured by CD154 and cytokine co-expression after in vitro stimulation with peanut
6. Change in frequency of peanut-specific T cells expressing CD137 and regulatory markers or producing IL-10 after in vitro stimulation with peanut
7. Change in peripheral blood transcriptional signature via RNAseq

Secondary mechanistic objectives

Secondary mechanistic objectives are to:

- Determine if immune and transcriptomic measures obtained early in the course of ingestion exposure can predict the development of desensitization or sustained unresponsiveness
- Identify immune and genomic biomarkers of the high threshold phenotype
- (**Exploratory**) Identify early-appearing functional pathways underlying a successful desensitization course

Secondary Mechanistic Endpoints

1. Immune and transcriptomic measures obtained early in the course of ingestion exposure that predict the development of desensitization or sustained unresponsiveness
2. Immune and genomic biomarkers of the high threshold phenotype
3. (**Exploratory**) Early-appearing functional pathways underlying a successful desensitization course.

Statistical Analyses

Immunocap Assay: Changes in serum component proteins measured through ImmunoCAP assay will be compared between randomization groups using linear mixed effects models for data collected at baseline, week 16 and the desensitization double-blind, placebo-controlled food challenge visit. The analysis will include a random subject effect and the fixed effect of randomization assignment, time, interaction between randomization assignment and time, and age. Spline models may be considered if the curves exhibit non-linear trends. Values below the limit of detection (LOD) will be imputed as X/sqrt(x)). The linear mixed effects model has the advantage that the estimation of the model parameters will be unbiased even in the presence of missing outcomes, assuming that the missing values depend only on the observed values (missing at random). If the missing at random assumption is not plausible, pattern-mixture modeling (which stratifies subjects by their pattern of missing data) will be used.

Cell population frequency: Quantile-linear mixed effects models will be used for the frequency of %CD63+ basophil, peanut-specific Th2 cells and T cells.

BBEA Assay: The bead-based epitope assay (BBEA) protocol will be used to obtain peanut specific of 64 15-mer epitopes derived from key peanut allergens—34 from Ara h 1, 16 from Ara h 2, and 14 from Ara h. All data processing, quality control, and analyses will be performed in R (version 4.3.3; R Foundation, Vienna, Austria). Binding scores for each epitope in the IgE and IgG4 assays will be obtained via *bbeaR*. Plate effect will be assessed using Principal Variation Component Analysis along with age strata, randomization assignment, visit, the interaction between randomization assignment and visit, and participant. If Plate effect is considered to be large, ie >10% contribution to the variance, Plate will be included as a factor in all analysis. Linear mixed effects models will be fitted by using the *limma* framework– for high dimensional omic data.

Transcriptomic analysis: mRNA libraries will be sequenced on the Illumina HiSeq 4000 System targeting 30 million 150 bp paired-end reads, according to the manufacturer's protocol. The sequencing reads will undergo quality control using FastQC and will be mapped to GRCh38 using STAR v2.4.0g1 aligner. Transcript counts will be estimated using featureCounts. The transcriptome profiles will be filtered to remove low expression genes with cpm≤1 in at least 10% of samples using edgeR2. The transcriptome data will be normalized using voom in the limma R package and scaled from -1 to 1. Linear mixed effects models will be fitted for each blood transcript and cell fraction using the lme4 R package. Sex, age, challenge (peanut or placebo), time point (0, 2, 4 hours), reaction threshold (0-5), and interactions between time,

challenge, and threshold will be included as fixed effects and a random intercept for each subject.