NOVARTIS

Clinical Development

QGE031/Ligelizumab

CQGE031G12301 / NCT04984876

A 52 week, multi-center, randomized, double-blind placebo-controlled study to assess the clinical efficacy and safety of ligelizumab (QGE031) in decreasing the sensitivity to peanuts in patients with peanut allergy

Statistical Analysis Plan (SAP)

Document type:	SAP Documentation Amendment 3
Release date:	11-September-2023
Number of pages:	48

Template Version 4.0, Effective from 23-Apr-2021

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
13 July 2021	Prior to DB lock	Creation of final version	N/A - First version	NA
28 July 2023	After Protocol Amendment 1 and before Dry run	Recruitment halt with the intention to terminate study early. Aligned with protocol amendment 1. Analyses changed.	Removed references to the second CSR	Section 1
			Described changes from early study termination and recruitment halt.	Section 1.1
			Removed analysis concerning sustainability of effect at Week 52.	Section 2.6.4
			Analysis of children under 12 years old removed, due to recruitment halt. Specific output to be presented at Primary Analysis listed. Analysis at Week 52 removed. DMC information added.	Section 2.13
			Due to halt of recruitment, sentences about children 6-11 years old removed.	Section 3
			[redacted]	[redacted]
			Handling of intercurrent events was simplified, with COVID-19 no longer being an intercurrent event and composite variable strategy being used consistently across endpoints. (Section 2.7.3 on 'Handling of Intercurrent Events' was deleted).	Section 1.2.1 and Section 1.2.2
			[redacted] the independent DMC is described.	[redacted]
			Treatment group labels are described.	Section 2.1.1.1
			The maximum tolerated dose is defined.	Section 2.1.1.5
			Described how mis-randomized patients are analyzed. PRO data collected removed and described elsewhere.	Section 2.2
			Potential subgroup analysis to split data according to notification of recruitment halt was added. Removed separate analysis for Japanese patients	Section 2.2.1
			Changes to how the number of patients will be summarized for patient disposition and included 12-week timepoint.	Section 2.3.1
			Added additional categories of baseline characteristics for summary.	Section 2.3.2

Added information about prior medications, concomitant medications, and epinephrine use. Replaced summaries by ATC classification by primary system organ class and MedDRA preferred term.	Section 2.4.2
Added log-transformed MTD (Maximum Tolerated Dose) at screening DBPCFC as covariate in logistic regression model. Added P-value will be generated with a likelihood ratio test.	Section 2.5.2
Removed tipping point analyses as missing data due to intercurrent events will be handled with relative conservative strategy (i.e., composite variable strategy)	Section 2.5.3
Added Kaplan-Meier plot and modified marginal standardization analysis.	Section 2.5.4
Maximum tolerated dose analyses and skin prick test described in more detail.	Section 2.6.4
Pooling periods for safety analyses described elsewhere (Section 2.1.1.1).	Section 2.7
Added details on treatment-emergent adverse events.	Section 2.7.1
Added exposure-adjusted incidence rate.	Section 2.7.1.1
Added details for summarizing injection site reactions.	Section 2.7.1.2
Imputation methods described for values that are abnormal or outside the level of quantification.	Section 2.7.3 and Section 2.8
Details on ECG parameters provided.	Section 2.7.4.1
[redacted].	[redacted]
IgE levels specified as at screening.	Section 2.9
Age groups and scoring of PRO instruments clarified, including SF-6Dv2.	Section 2.10.1
Analysis for PRO scores described.	Section 2.10.1.1
[redacted].	[redacted]
Power calculations updated using PASS Version 11. Paragraph on consistency trend across age groups removed.	Section 3.1
Sentence about potentially increasing sample size for covid impact was removed. Power table for children 6-11 removed.	Section 3.2
Table on precision levels for safety incidence rate per age group removed.	Section 3.2.1
Changes to protocol specified analyses noted.	Section 4
Details for multiple imputation added, for maximum tolerated dose, binary endpoints involving logistic regression, and marginal standardization. Tipping point analysis for primary estimand removed.	Section 5.4.1
For maximum severity of symptoms, covariates were specified and details on imputation strategy added.	Section 5.4.2
IgE specified as being measured at screening.	Section 5.4.2.1

11 Sep 2023	Before database lock	Clarification of analyses	Specified calculation of negative control for skin prick test mean diameters.	Section 2.6.4
			Removed SF-36 output from Table 2-2 and specified missing data approach for additional analysis	Section 2.13
			Labelled systolic and diastolic blood pressure values	Section 5.3

Table of contents

List of tables

List of figures

List of abbreviations

[redacted]	[redacted]
AE	Adverse Event
ANCOVA	Analysis of covariance
ATC	Anatomical Therapeutic Chemical
AUC	Area under Curve
bpm	Beats per minute
CD23	IgE (low affinity) receptor
CI	confidence intervals
CM	Concomitant Medication
Covid-19	Corona Virus Disease 2019
CRF	Case Report/Record Form (paper or electronic)
CSR	Clinical study report
CV	Coefficient of variation
DBL	Database Lock
DBPCFC	Double Blind Placebo Controlled Food Challenge
DMC	Data Monitoring Committee
[redacted]	[redacted]
ECG	Electrocardiogram
eCRF	Electronic case report form
EOS	End of study
FAIM	Food Allergy Independent Measure
FAQLQ	Food Allergy Quality of Life Questionnaire
FAS	Full Analysis Set
FcεRI	IgE (high affinity) receptor
gMCP	Graph Based Multiple Comparison Procedures
h	Hour
HRQoL	Health-Related Quality of Life
IgE	Immunoglobulin E *(specific and un-specific for the study allergen)*
IgG	Immunoglobulin G
L	Liter
LLOQ	Lower limit of quantification
MAR	Missing at random
MedDRA	Medical dictionary for regulatory activities
mg	Milligram(s)
mL	Milliliter(s)
mm	Millimeter
mmHg	Blood pressure unit millimeters of mercury
MRI	Magnetic Resonance Imaging
ms	Millisecond
MTD	Maximum Tolerated Dose
OFC	Oral Food Challenge
[redacted]	[redacted]
[redacted]	[redacted]
PRO	Patient Reported Outcomes
q4W	Every 4 weeks

RAS	Randomized Analysis Set
RR	Responder rate
SABA	Short-acting ß-agonist
SAE	Serious Adverse Event
SAF	Safety analysis set
SC	Subcutaneous
SCq4W	Subcutaneous injection every 4 weeks
SD	Standard deviation
sIgE	Specific IgE for the study allergen
SMQ	Standardized MedDRA Query
SPT	Skin Prick Test
Total IgE	Drug-bound and free IgE
µL	Microliter
ULOQ	Upper limit of quantification

1 Introduction

The purpose of this document is to describe the statistical analyses mentioned in version 01 (amendment 1) of the Phase 3 study CQGE031G12301 protocol.

This document covers statistical and analytical plans for the primary endpoint analysis (i.e., after all participants have completed the visit at Week 12) and the final analysis (i.e., after all participants have completed the follow-up visit or prematurely withdrawn from the study).

1.1 Study design

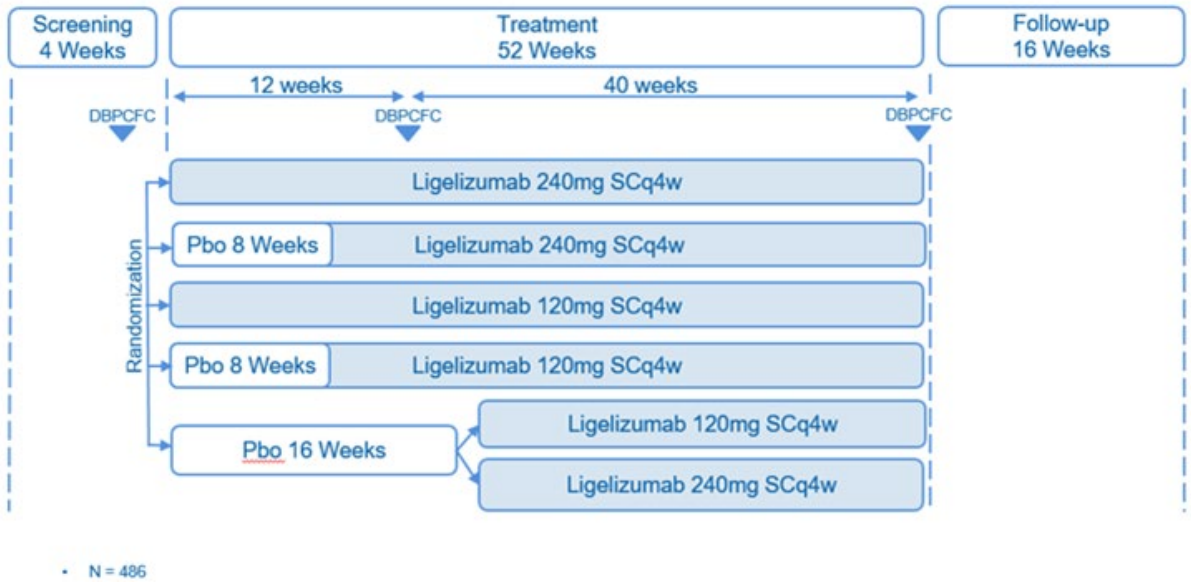
This is a 52-week, Phase 3 multi-center, randomized, double-blind and placebo-controlled study to assess the safety and clinical efficacy of ligelizumab (QGE031) in participants with a medically confirmed diagnosis of IgE-mediated peanut allergy. The primary objective of the study is to evaluate the efficacy of ligelizumab 240 mg and 120 mg (SCq4w) compared to placebo at Week 12, as measured by the proportion of participants who can tolerate a single dose of ≥ 600 mg (1044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the double-blind, placebo controlled, food challenge (DBPCFC).

A total sample size of approximately 486 randomized participants aged 6 to 55 years was initially planned. Due to the decision of halt recruitment and terminate the study early (Section 4), the final size is 211, with only adolescent and adult participants being recruited. Participants were randomized at a 2:2:2:2:1 ratio to one of five treatment arms at the time of randomization visit as follows:

- Ligelizumab 240 mg SC q4w
- Ligelizumab 120 mg SC q4w
- Placebo 8 weeks to Ligelizumab 240 mg SC q4w
- Placebo 8 weeks to Ligelizumab 120 mg SC q4w
- Placebo 16 weeks to Ligelizumab 240 or 120 mg SC q4w (with a ratio of 0.5 : 0.5 for switched treatment)

Participants will be stratified based on region, total IgE at screening (<350 IU/ mL; ≥350 IU/ mL) and age (12 - 17y, and 18 - 55y).

Figure 1-1 Study design



The study will include the following:

1. **Screening period (Duration of 4 weeks)**

2. **Treatment period (Duration of 52 weeks)**
   - **Double-blind placebo-controlled** treatment period **(Duration of 12 weeks):** Study participants will be seen in the clinic at Day 1, Week 1, Week 4, Week 8 and Week 12. The DBPCFC will be performed at Week 12. Participants assigned to the 8-week placebo arms will receive the first dose of blinded ligelizumab treatment at the Week 8 visit. Participants assigned to the 16-week placebo arm will receive the first dose of blinded ligelizumab treatment at the Week 16 visit.
   - **Long-term active blinded (no placebo control)** treatment period **(Duration of 40 weeks/ 36 weeks for participants on the 16-week placebo arm):** Starting at Week 16, all study participants will receive blinded ligelizumab study treatment in the clinic every 4 weeks until Week 52. A final DBPCFC will be performed at Week 52.

3. **Post-treatment follow-up period (Duration of 16 weeks):** There are 4 planned clinic visits (every 4 weeks). Study evaluations include safety, [redacted], post-treatment biomarkers. Study treatment is not given and there is no DBPCFC.

At the start of the study, recruitment was restricted to 12 - 55 year old participants. When approximately 60 adolescent participants (defined as 12 -17 years of age) would have completed all Week 12 assessments, an interim analysis [redacted] was to be performed for confirmation of the pediatric dose (safety will be reviewed by a Data Monitoring Committee - DMC). Independent sponsor members who are responsible for [redacted] Modeling & Simulation will be unblinded to the results of this interim analysis.

As Novartis decided to halt recruitment (Date of issuing halt notification: 05 Apr 2023) with the intention to terminate the study early, there will be no recruitment of children under age 12.

Patients who are already in the study (i.e., signed informed consent) at the time of the halt notification will be managed according to their individual progress within the study as described in the table below.

Group 1	*Already completed Week 52 Visit assessments*	Complete the 16-week safety follow up period as per protocol
Group 2	*Already completed Week 12 Part 2 Visit*	Attend the next scheduled visit (as per schedule of events), which will now correspond to the end of treatment visit (Week 52 Part 1). Patients will be informed at that visit that the treatment is stopped (no drug administration at that visit) and that they will enter the 16-week follow up period as per protocol. Notably, no oral food challenge will be done at this end of treatment visit
Group 3	*Randomized but have not yet reached the Week 12 Part 2 Visit*	At their next visit, investigators should inform the patient of the study status and provide them with the opportunity to continue and perform the assessments until week 12 Part 2. The week 12 Part 2 visit will now be considered end of treatment (Week 52 Part 2) so that afterwards patients will then enter the 16-week safety follow up period as per protocol. In this situation, there is no need to fill out the "Trial Feedback Questionnaire."
Group 4	*Reached the Screening Visit 2 Part 1*	Allowed to complete the Screening Visit 2 Part 2 (i.e., complete both parts of the baseline Oral Food Challenge) and, pending final eligibility, to be randomized and continue as patients in Group 3.
Group 5	*In screening period but have not reached the Screening Visit 2 Part 1*	Screening failed and patients will be **eligible for future pivotal studies**

1.2 Study objectives, endpoints and estimands

Table 1-1 Objectives and related endpoints

Objective(s)	**Endpoint(s)**
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• To evaluate the efficacy of ligelizumab 240 mg and 120 mg (SCq4w), compared to placebo, as measured by the proportion of participants who can tolerate a single dose of ≥ 600 mg (1044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC at Week 12	• Responder status defined as tolerating a single dose of ≥ 600 mg (1044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC conducted at Week 12.
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
• **Key secondary objectives** • To evaluate the efficacy of ligelizumab 240mg and 120mg (SCq4w), compared to placebo, as measured by the proportion of participants who can tolerate a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC at Week 12	• **Key secondary endpoints** • Responder status defined as tolerating a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC conducted at Week 12.
• To evaluate the efficacy of ligelizumab 240mg and 120mg (SCq4w), compared to placebo, as	• Responder status defined as tolerating a single dose of 3000 mg (5044 mg cumulative tolerated

Objective(s)	Endpoint(s)
measured by the proportion of participants who can tolerate a single dose of 3000 mg (5044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC at Week 12	dose) of peanut protein without dose-limiting symptoms during the DBPCFC conducted at Week 12.
• To evaluate the efficacy of ligelizumab 240mg and 120mg (SCq4w), compared to placebo, as measured by the maximum symptom severity at any single challenge dose up to and including 1000 mg of peanut protein during the DBPCFC at Week 12	• Maximum severity of symptoms occurring at any challenge dose of peanut protein up to and including 1000 mg during the DBPCFC conducted at Week 12. Symptom severity will be categorized as 4 levels: None, Mild, Moderate, Severe.
• To evaluate the efficacy of 8 weeks of placebo treatment followed by 4 weeks of ligelizumab 120 mg / 240 mg (SCq4w) treatment compared to 12 weeks of placebo treatment, as measured by the proportion of participants who can tolerate a single dose ≥1000 mg of peanut protein without dose-limiting symptoms during the DBPCFC at Week 12	• Responder status defined as participants tolerating a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC conducted at Week 12 (8 weeks of placebo + 4 weeks of ligelizumab treatment vs. 12 weeks of placebo).
• **Other secondary objectives**	• **Other secondary endpoints**
• To evaluate the efficacy of ligelizumab 240mg and 120mg (SCq4w), as measured by the proportion of participants who can tolerate a single dose of ≥1000 mg (2044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC at Week 52 compared to Week 12	• Proportion of participants tolerating a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during DBPCFC conducted at Week 52 • Change in maximum tolerated dose (MTD) of peanut protein without dose-limiting symptoms during the DBPCFC at Week 52 compared to Week 12
• To evaluate the effects of ligelizumab 240 mg and 120 mg (SCq4w), compared to placebo (when applicable), as measured by multiple systemic biomarkers to inform on response to treatment or disease severity	• Change from baseline at Week 12, 16 and Week 52 of •peanut-specific IgE •peanut-specific IgG4
• To evaluate the safety and tolerability of ligelizumab 240 mg and 120 mg (SCq4w)	• Summaries of treatment-emergent adverse events, vital signs, ECG, and laboratory values
• To assess the ability of IgE suppression to impact skin mast cells through the assessment of allergen-specific skin prick test (SPT).	• Change from baseline (screening) in SPT mean wheal diameters at Week 16, Week 56 and Week 68.
• To evaluate the impact of ligelizumab on the health-related quality of life (HRQoL) of patients with peanut allergy.	• Change from baseline in total and domain scores in the FAQLQ, FAIM, and SF-36v2 by age and responder (subject and/or caregiver) at various points in time.

[redacted]

Objective(s)	Endpoint(s)
[redacted]	

1.2.1 Primary estimand(s)

The primary estimand quantifies the effects of ligelizumab 120mg and 240mg SCq4w as compared to placebo on the proportion of responders at Week 12 regardless of intake of rescue medication prior to Week 12.

The primary estimand is described by the following attributes:

- **Population**: participants who have been diagnosed with IgE-mediated peanut allergy and met study inclusion/exclusion criteria.
- **Variable**: Responder status defined as tolerating a single dose of ≥ 600 mg (1044 mg cumulative tolerated dose*) of peanut protein without dose-limiting symptoms **during the DBPCFC conducted at Week 12.
- **Treatment**: the randomized treatment^ (ligelizumab 120mg SCq4w, ligelizumab 240mg SCq4w, and placebo) plus rescue medication (e.g., epinephrine, SABA, anti-histamines), if needed.
- **Handling of intercurrent events:**
  - Discontinuation of treatment or missing more than one dose prior to Week 12: participants who discontinue treatment or miss more than one dose prior to Week 12 will not undergo any DBPCFC. They will be considered non-responders (*composite variable strategy*).
  - Intake of rescue medication prior to DBPCFC conducted at Week 12: ignorable (*treatment policy strategy*, reflected in the Treatment attribute)

    Epinephrine, SABA, anti-histamines and saline bolus are typically used as rescue medication to treat allergic reactions. Participants who experience allergic reactions due to accidental exposure or other reasons may need to take rescue medications prior to Week 12. However, participants must be in good health before proceeding with the food challenge and meet the washout requirement of permitted medications as described in Protocol Table 6-2. Therefore, to reflect clinical practice, all observed values after this intercurrent event will be used in the statistical analyses whenever the treatment effect is estimated.

- **Summary measure**: odds ratio comparing the proportion of responders between each ligelizumab dose group and placebo group
- **The cumulative tolerated dose is the sum of the tolerated doses, not including the reactive dose* (Casale et al 2019)
- ***Dose-limiting symptoms indicate a true allergic reaction occurring during administration of a single dose of peanut protein at the DBPCFC that should preclude the administration of any further doses in the view of the investigator. Symptoms that require administration of any rescue medication are considered dose-limiting symptoms.*
- *^ The randomized treatment indicates ligelizumab 120 mg arm, ligelizumab 240 mg arm, and placebo 16wk to ligelizumab 120/240 mg arm (labelled as placebo for efficacy estimands).*

1.2.2 Secondary estimand(s)

The four key secondary endpoints defined in the protocol are considered for estimands as below.

1.2.2.1 Secondary estimand 1 : Proportion of responders who can tolerate a single dose of ≥1000 mg of peanut protein at Week 12

The secondary estimand 1 is the same as the primary estimand except that "a single dose of ≥ 600 mg (1044 mg cumulative tolerated dose)" is replaced by "a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose)".

1.2.2.2 Secondary estimand 2: Proportion of responders who can tolerate a single dose of 3000 mg of peanut protein at Week 12

The secondary estimand 2 is the same as the primary estimand except that "a single dose of ≥ 600 mg (1044 mg cumulative tolerated dose)" is replaced by "a single dose of ≥ 3000 mg (5044 mg cumulative tolerated dose)".

1.2.2.3 Secondary estimand 3: Maximum severity of symptoms up to and including 1000 mg of peanut protein

This secondary estimand 3 is described by the following attributes:

- **Population**: same as for the primary estimand.
- **Variable**: maximum severity of symptoms occurring at any challenge dose of peanut protein up to and including 1000 mg during the DBPCFC conducted at Week 12. According to CoFAR grading scale of dose-limiting symptoms, symptom severity will be categorized as Mild, Moderate and Severe. Symptom severity for participants who completed DBPCFC without any symptoms will be categorized as "None".
- **Treatment**: same as for the primary estimand.
- **Handling of intercurrent events:**
  - Discontinuation of treatment or missing more than one dose prior to Week 12: Missing data at Week 12 will replaced by the maximum severity of symptoms at screening (*composite variable strategy*).

- Intake of rescue medication prior to DBPCFC conducted at Week 12: ignorable (*treatment policy strategy*, reflected in the "Treatment" attribute).

- **Summary measure**: odds ratio comparing the odds of developing less severe symptoms between each ligelizumab dose group and placebo group.

1.2.2.4 Secondary estimand 4: Proportion of responders who can tolerate a single dose of 1000 mg of peanut protein at Week 12 with 8 weeks of placebo followed by 4 weeks of ligelizumab

The secondary estimand 4 provides insight into the treatment effect after one dose of ligelizumab. This secondary estimand is the same as the primary estimand except the following attributes :

- **Variable**: Responder status defined as tolerating a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC conducted at Week 12.
- **Treatment**: the randomized treatment (placebo 8 weeks to ligelizumab 240 mg SCq4w, placebo 8 weeks to ligelizumab 120 mg SCq4w and placebo 16 weeks to ligelizumab 240/120 mg SCq4w) plus rescue medication, if needed.

2 Statistical methods

2.1 Data analysis general information

The statistical analysis of the study will be performed by [redacted] a designated Contract Research Organization (CRO) following this document. SAS version 9.4 will be used to perform all analyses.

There are two planned analyses before the final database lock, in addition to the DMC analyses. The details of these analyses are provided in the Section 2.13 below.

An independent DMC will conduct periodic monitoring of safety data and emerging risk/benefit. Interim reports to the DMC will be generated by an independent statistical group not involved in the conduct of the trial. Statistical Analysis Plan for the DMC analyses will be prepared separately.

Categorical data will be summarized as frequencies and percentages. For continuous data, mean (geometric mean for non-normal distributed variables), standard deviation, median, 25th and 75th percentiles, minimum and maximum will be presented.

2.1.1 General definitions

2.1.1.1 Study treatment

The study treatment includes investigational drug QGE031 and placebo. The following abbreviated treatment groups will be used as the headers in the tables.

- QGE031 240 mg q4w
- QGE031 120 mg q4w
- Placebo 8 wks – QGE031 240 mg

- Placebo 8 wks – QGE031 120 mg
- Placebo 16 wks – QGE120 or 240 mg q4w*

* In the efficacy analysis outputs, placebo 16 wks switch will be labelled as "placebo group" for simplicity.

In addition, for most of safety summaries, the following abbreviated treatment groups under different pooling period will be used as the headers in the tables.

- Placebo-controlled pool up to Week 8 (QGE031 240 mg vs. QGE031 120 mg vs. three Placebo arms combined)
- Placebo-controlled pool up to Week 16 (QGE031 240 mg vs. QGE031 120 mg vs. Placebo 16 wks – QGE120 or 240 mg)
- Entire study period up to Week 68 (combined three arms with QGE031 240 mg vs. combined three arms with QGE 120 mg)

2.1.1.2 Study day

Study day will be defined as the number of days since the date of first dose of study treatment. The date of first dose of study treatment will be defined as Day 1 and the day before the first dose of study treatment will be defined as Day -1.

Therefore, for a particular date, study day will be calculated as follows:

For dates on or after the first dose date of study treatment, study day will be calculated as [Assessment date – Date of first dose of study treatment + 1]. For dates prior to the first dose date of study treatment, study day will be calculated as [Assessment date – Date of first dose of study treatment].

2.1.1.3 Baseline definition

In general, baseline is defined as the last measurement before the first dose of study treatment. Details on calculation of baseline for PRO measures is provided in Section 2.10.

2.1.1.4 Post-baseline definition

Post-baseline measurements are defined as the assessments after the first dose of study treatment. When change from baseline of raw or log-transformed data is of interest, the following formula will be used to calculate the change from baseline, provided, both baseline and post-baseline values are available:

Change from baseline = post-baseline value – baseline value.

2.1.1.5 Maximum tolerated dose in DBPCFC

Maximum tolerated dose (MTD) is defined as the maximum tolerated single dose without any dose limiting symptoms in peanut challenge part. If the MTD at any visit is 0mg, it will be imputed as 0.3 mg and reported as "< 1 mg."

2.2 Analysis sets

The Randomized Analysis Set (RAS) consists of all randomized participants, regardless of whether or not they receive a dose of study drug. Participants will be analyzed according to the treatment they are assigned to.

The Full Analysis Set (FAS) comprises all participants who received at least one dose of study treatment. According to the intent to treat principle, participants will be analyzed according to the treatment they have been assigned to during the randomization procedure. Mis-randomized patients (mis-randomized in IRT) will be excluded. Mis-randomized patients are defined as cases where IRT contacts were made by the site either prematurely or inappropriately prior to confirmation of the patient's final randomization eligibility and no study medication was administered to the patient. FAS will be used for all efficacy variables, unless otherwise stated.

The Safety Set (SAF) includes all participants who received at least one dose of study treatment whether or not being randomized. Participants will be analyzed according to the treatment they received, where treatment received is defined as the randomized/assigned treatment if the participant took at least one dose of that treatment or the first treatment received if the randomized/assigned treatment was never received. The safety set will be used in the analysis of all safety variables.

[redacted]

2.2.1 Subgroups of interest

Subgroup analyses will be performed for primary endpoint and selected secondary endpoints using the same models mentioned in Section 2.6 and Section 2.7 but with additional model terms for the subgroup (if not already included in the model) and subgroup-by-treatment interaction terms. Due to the reduced number of participants who completed the study, models for subgroup analyses may fail to converge after dropping factors, then only descriptive subgroup analyses will be provided instead. The subgroup variables are listed below:

- By age group (12 – 17 years, 18 – 55 years)
- By total IgE at screening (<350 IU/ mL; ≥350 IU/ mL)
- By age group and total IgE at screening
- By maximum tolerated dose of allergen at screening OFC (< 30 mg; ≥ 30 mg)
- By allergy status (mono- vs poly-sensitized participants)

Mono-sensitized participants are only sensitized to peanut protein. Poly-sensitized participants are sensitized to peanut protein and at least one of the other proteins in the panel. Poly-sensitization is defined as sIgE ≥ 0.35 kUA/L for a food allergen in the panel other than peanut and confirmed in the medical history.

The following subgroup analysis may be conducted to evaluate potential impact of temporary halt on primary and key secondary efficacy endpoints, if participants who have not reached

Week 12 upon receiving the notification of temporary halt (Date of issuing temporary halt notification : 05 Apr 2023) have a dropout rate greater than or equal to 25%:

- By Week 12 completion date (or date of study discontinuation if study discontinuation occurred prior to Week 12) before or after the notification of the halt.

In addition, subgroup analyses by age (12-17 years and 18-55 years) for selected safety endpoints, [redacted] will also be provided.

2.3 Patient disposition, demographics, and other baseline characteristics

2.3.1 Patient disposition

For each study epoch (i.e., screening, treatment phase and post treatment follow-up), the overall number of participants who entered, completed and discontinued that phase will be summarized including the primary reasons for discontinuation of that phase.

The number of participants who completed the 12-week and 52-week study treatment and who discontinued prematurely will be shown including the reasons for discontinuation of treatment.

Number of participants with protocol deviations will be summarized by category and deviation.

2.3.2 Demographics and other baseline characteristics

Demographic and other baseline data including disease characteristics will be summarized descriptively by treatment group for the RAS.

Demographics

- Age
- Age group (12-17 years and 18-55 years)
- Sex
- Race
- Ethnicity
- Region [Japan; US and Canada (North America); Australia; Denmark, France, Germany, Italy, Spain, the Netherlands (EU)]
- Weight
- Height
- Body Mass Index – calculated as weight (kg)/(height (m))^2
- BMI group (< 25, 25 - < 30, ≥ 30 kg/ m^2)

Baseline characteristics

- Peanut sIgE
- Peanut specific IgG4
- SPT mean wheal diameter
- MTD (maximum tolerated dose) of peanut protein at screening DBPCFC (<1< 1, 1, 3, 10, 30 mg)

- Poly-allergic to food
- History of anaphylactic reaction to food (Yes, No)
- Number of Use of Epinephrine auto-injector in the last two years
- Duration of peanut allergy (months since peanut allergy diagnosis)
- History of desensitization therapy to peanut allergy (Yes, No)

Categorical data will be presented as frequencies and percentages. For continuous data, mean (geometric mean for non-normal distributed variables), standard deviation, median, minimum, and maximum will be presented.

Relevant medical histories and current medical conditions at baseline will be summarized by system organ class and preferred term, by treatment group.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

The duration of exposure in weeks to each treatment group will be summarized using descriptive statistics.

Duration of exposure to study treatment will be calculated as the number of weeks between the first dose date and the last dose date exposed to that treatment over the specified period (Duration of exposure = (date of last known study treatment – date of first known study treatment + 28)/7).

In addition, the number of doses, total cumulative dose and number of missed doses will be presented. Categorical data will be summarized as frequencies and percentages. For continuous data, mean, standard deviation, median, 25^{th} and 75^{th} percentiles, minimum and maximum will be presented.

2.4.2 Prior, concomitant and post therapies

Prior medications are defined as treatment taken and stopped prior to first dose or study treatment. Prior medications will be summarized by Anatomical Therapeutic Chemical (ATC) code, preferred term and treatment group on RAN set, separately for peanut allergy treatment (other than the ones taken during screening DBPCFC) and non-peanut allergy treatment (including ones taken during screening DBPCFC). For peanut allergy treatment, reason for discontinuation will be summarized as well.

Concomitant medications are defined as any medication given at least once between the day of first dose of study treatment and the date of the last study visit. Concomitant medications will be summarized by ATC code, preferred term and treatment group for safety set.

Rescue medications will be summarized by ATC code, preferred term and treatment group for safety set.

Use of epinephrine during and outside of OFC will be summarized separately. Epinephrine used during the OFC are reported by the Investigator, on the concomitant medication eCRF page

after choosing Yes to the question on OFC signs and symptoms eCRF page: Was any treatment provided for these sign/symptoms?

Significant prior and concomitant surgeries and procedures will be summarized by primary system organ class and MedDRA preferred term.

2.5 Analysis supporting primary objective(s)

This section will describe the statistical analysis of the primary estimand. Details of the hypothesis testing strategy including primary and key secondary endpoints to handle multiplicity are provided in Section 2.6.2.

2.5.1 Primary endpoint(s)

Definition of primary estimand is provided in Section 1.2.1.

2.5.2 Statistical hypothesis, model, and method of analysis

The trial will be considered positive, if at least one of the two ligelizumab doses demonstrates a statistically significant result in comparing the proportion of responders as described below. The two doses will be tested in parallel for 12 weeks treatment on the primary endpoint.

Let p_j denote the responder rate for treatment regimens j, j=0, 1 or 2 where

- 0 corresponds to placebo
- 1 corresponds to ligelizumab 120 mg treated for 12 weeks
- 2 corresponds to ligelizumab 240 mg treated for 12 weeks

Responder rate is defined as the proportion of participants tolerating a single dose of ≥ 600 mg (1044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during the DBPCFC conducted at Week 12. For participants with treatment discontinuation or missing more than 1 doses of study drug prior to Week 12 due to reasons other than operational complications caused by public health emergency, they will be considered non-responders.

The following hypotheses will be tested:

$H_{0\ 120\ RR600}: p_1 / (1- p_1) \leq p_0 / (1- p_0)$ versus $H_{A\ 120\ RR600}: p_1 / (1- p_1) > p_0 / (1- p_0)$

$H_{0\ 240\ RR600}: p_2 / (1- p_2) \leq p_0 / (1- p_0)$, $H_{A\ 240\ RR600}: p_2 / (1- p_2) > p_0 / (1- p_0)$

$H_{0\ 120\ RR600}$: ligelizumab 120mg is not superior to placebo with respect to the responder rate at a level of 600 mg peanut protein (1044 mg cumulative tolerated dose) without dose-limiting symptoms at Week 12

$H_{0\ 240\ RR600}$:ligelizumab 240mg is not superior to placebo with respect to the responder rate at a level of 600 mg peanut protein (cumulative tolerated dose 1044 mg) without dose-limiting symptoms at Week 12

The primary endpoint will be analyzed based on a logistic regression model, including treatment, age subgroup (12 – 17 years, 18 – 55 years), region as fixed class effects, as well as log-transformed total IgE at screening and log-transformed MTD (Maximum Tolerated Dose) at screening DBPCFC as covariates. P-value will be generated with a likelihood ratio test. Odds

ratio and 95% profile likelihood confidence intervals (CI) will be presented comparing each ligelizumab dose to placebo with respect to the proportions of responders.

The detailed testing strategy for primary and key secondary endpoints is provided in Section 2.6.2.

2.5.3 Handling of missing values not related to intercurrent event

Handling of intercurrent events associated with key secondary objectives are described in the Section 1.2.1. Missing data not related to intercurrent events will be imputed for primary endpoint using multiple imputation under the missing at random (MAR) assumption. Details of the imputation method are mentioned in the Section 5.4.1.

2.5.4 Supplementary analyses

Two supplementary analyses will be performed for the primary endpoint.

Marginal Standardization

The first supplementary estimand is the same as for the primary estimand except that the summary measure is the risk difference for marginal proportions of responders . The risk difference will be derived from the predicted risks for every patient as if they had received each treatment using a logistic regression model (Ge et al 2011). The model will include the same terms as in the primary analysis. The estimated treatment group risks, risk difference and 95% CI will be presented. Details are provided in Section 5.4.1.

Kaplan-Meier plot

A Kaplan-Meier plot will show the probability of experiencing a dose-limiting symptom during dose escalation of DBPCFC at Week 12 by treatment group. The dose of peanut protein at which dose-limiting symptom occurs during the DBPCFC at Week 12 will be treated as the time variable. There will be two separate plots showing those who completed DBPCFC at Week 12: the first will show the results for the overall population; the second will be stratified by the highest tolerated dose of peanut protein at baseline (< 30 mg or >= 30mg). Participants' maximum tolerated dose may be censored if they:

- tolerated 3000mg of peanut protein; or
- completed a certain level of challenge dose without experiencing dose-limiting-symptoms and chose to stop the DBPCFC before beginning or finishing a higher challenge dose (for example, due to the taste of the challenge material).

2.6 Analysis supporting secondary objectives

This section describes the analyses of key secondary and other secondary endpoints. The details of the analyses are described below.

2.6.1 Key secondary endpoint(s)

Definition of key secondary endpoints is provided in Section 1.2.2.

The other secondary endpoints and corresponding analyses are defined in the Section 2.6.5 below.

2.6.2 Statistical hypothesis, model, and method of analysis

Key secondary endpoint analyses

The key secondary null hypotheses are as defined below. The familywise type I error rate will be controlled at the one-sided 0.025 level across the primary and key secondary null hypotheses in a closed testing procedure (Bretz et al 2009).

$H_{0\ 120\ RR1000}$: ligelizumab 120 mg is not superior to placebo with respect to the responder rate at a level of 1000 mg peanut protein (cumulative tolerated dose 2044 mg) at Week 12

$H_{0\ 240\ RR1000}$: ligelizumab 240 mg is not superior to placebo with respect to the responder rate at a level of 1000 mg peanut protein (cumulative tolerated dose 2044 mg) at Week 12

$H_{0\ 120\ RR3000}$: ligelizumab 120 mg is not superior to placebo with respect to the responder rate at a level of 3000 mg peanut protein (cumulative tolerated dose 5044 mg) at Week 12

$H_{0\ 240\ RR3000}$: ligelizumab 240 mg is not superior to placebo with respect to the responder rate at a level of 3000 mg peanut protein (cumulative tolerated dose 5044 mg) at Week 12

$H_{0\ 120\ severity}$: ligelizumab 120 mg is not superior to placebo with respect to the odds of developing less severe level of symptoms evaluated by maximum severity of symptoms at any challenge dose of peanut protein up to and including 1000 mg during the DBPCFC conducted at Week 12

$H_{0\ 240\ severity}$: ligelizumab 240 mg is not superior to placebo with respect to the odds of developing less severe level of symptoms evaluated by maximum severity of symptoms at any challenge dose of peanut protein up to and including 1000 mg during the DBPCFC conducted at Week 12

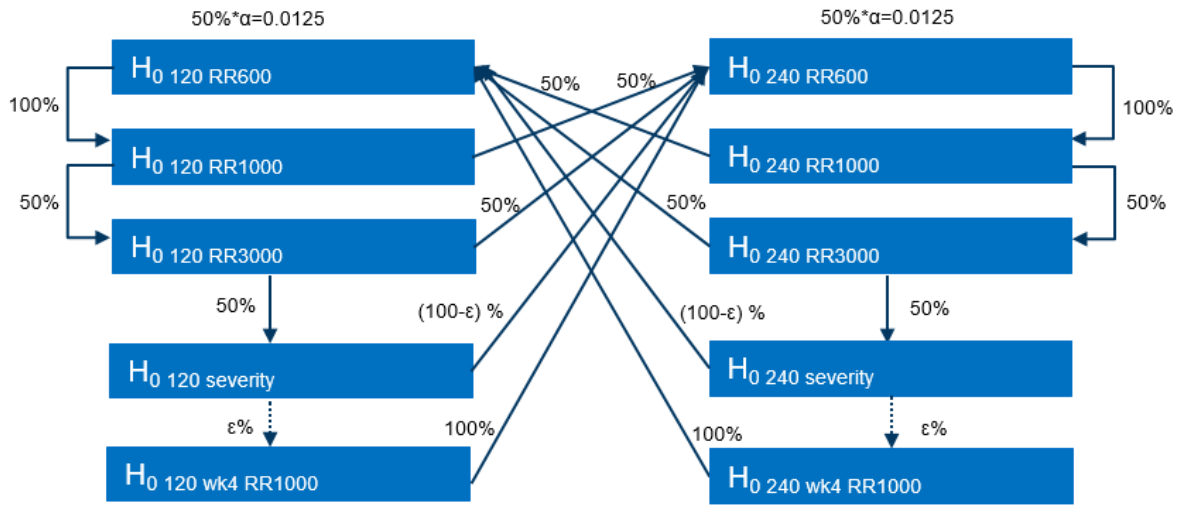
The following secondary hypotheses will be used to evaluate onset of action of ligelizumab (8 weeks of placebo + 4 weeks of ligelizumab treatment):

$H_{0\ 120\ wk4\ RR1000}$: ligelizumab 120 mg (8 weeks of placebo + 4 weeks of ligelizumab treatment) is not superior to placebo with respect to the responder rate at a level of 1000 mg peanut protein (cumulative tolerated dose 2044 mg) at Week 12

$H_{0\ 240\ wk4\ RR1000}$: ligelizumab 240 mg (8 weeks of placebo + 4 weeks of ligelizumab treatment) is not superior to placebo with respect to the responder rate at a level of 1000 mg peanut protein (cumulative tolerated dose 2044 mg) at Week 12

The graphical approach of Bretz et al 2009 for sequentially rejecting testing procedures is used to illustrate the testing strategy in Figure 2-1 below.

Figure 2-1 Testing strategy



- The hypotheses will be tested in two branches constituting ligelizumab 120 mg versus placebo (left branch) and ligelizumab 240 mg versus placebo (right branch). Initially, the full alpha level of 0.025 (one-sided) is equally split across the primary hypotheses for the two branches.
- Once this first primary null hypothesis $H_{0\ 120\ RR600}$ or $H_{0\ 240\ RR600}$ for a dose has been rejected at the initial alpha level of 0.0125 (one-sided), the alpha will be fully distributed to $H_{0\ 120\ RR1000}$ or $H_{0\ 240\ RR1000}$ for the same dose.
- If $H_{0\ 120\ RR1000}$ or $H_{0\ 240\ RR1000}$ is rejected for a dose, then 50% of its local significance level is reassigned to the primary null hypothesis for the other dose, and 50% of its local significance level is reassigned to $H_{0\ 120\ RR3000}$ or $H_{0\ 240\ RR3000}$ for the same dose.
- If $H_{0\ 120\ RR3000}$ or $H_{0\ 240\ RR3000}$ is rejected for a dose, then 50% of its local significance level is reassigned to the primary null hypothesis for the other dose, and 50% of its local significance level is reassigned to $H_{0\ 120\ severity}$ or $H_{0\ 240\ severity}$ for the same dose.
- If $H_{0\ 120\ severity}$ or $H_{0\ 240\ severity}$ is rejected for a dose, (100-ε)% of its local significance level is reassigned to the primary null hypothesis for the other dose. ε is set to a very small number in practice, e.g., 10^{-10}. The dotted dashed edges with a weight of ε indicate the local significance level will only be reassigned to $H_{0\ 120\ wk4\ RR1000}$ or $H_{0\ 240\ wk4\ RR1000}$, once both $H_{0\ 120\ severity}$ and $H_{0\ 240\ severity}$ are rejected.
- If $H_{0\ 120\ wk4\ RR1000}$ or $H_{0\ 240\ wk4\ RR1000}$ is rejected for a dose, 100% of its local significance level is reassigned to the primary null hypotheses for the other dose.

The proportion of responders for the endpoints as stated below will be analyzed using a logistic regression model in a similar fashion as the primary estimand (see Section 5.4.1 for details).

- Responder status defined as tolerating a single dose of 1000 mg (2044 mg cumulative tolerated dose) peanut protein without dose-limiting symptoms as Week 12.
- Responder status defined as tolerating a single dose of 3000 mg (5044 mg cumulative tolerated dose) peanut protein without dose-limiting symptoms at Week 12.

- Responder status defined as participants tolerating a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose) peanut protein without dose-limiting symptoms at Week 12 (8 weeks of placebo + 4 weeks of ligelizumab treatment vs. 12 weeks of placebo).

The odds of developing less severe level of symptoms for the endpoint, defined as maximum severity of symptoms occurring at any challenge dose of peanut protein up to and including 1000 mg during the DBPCFC conducted at Week 12, will be analyzed with a proportional odds model, including treatment, age subgroup (6-11 years, 12-17 years, 18-55 years), region as fixed class effects and log-transformed baseline total IgE at screening as a covariate.

2.6.3 Handling of missing values not related to intercurrent events

Handling of intercurrent events associated with key secondary objectives are described in the Section 1.2.2. The missing data not related to intercurrent events for key secondary will be imputed based on the missing at random (MAR) assumption. The details of the imputation method are mentioned in Section 5.4.1.

2.6.4 Other secondary endpoint analyses

Secondary endpoints which are not part of the testing strategy include:

- Proportion of participants tolerating a single dose of ≥ 1000 mg (2044 mg cumulative tolerated dose) of peanut protein without dose-limiting symptoms during DBPCFC conducted at Week 52

Number and percentage of responders at Week 52 will be summarized by treatment group in overall population. Number and percentage of responders at Week 52 may also be summarized by treatment group in the subset of responders at Week 12. Missing data of DBPCFC at Week 52 will not be imputed. The summary statistics will be provided for observed data.

- Change in maximum tolerated dose (MTD) of peanut protein without dose-limiting symptoms during the DBPCFC at Week 52 compared to Week 12

MTD at Week 52 along with changes from baseline and changes from Week 12 will be summarized descriptively. Missing data will not be imputed. The summary statistics will be provided for observed data.

- Change from baseline in MTD of peanut protein without dose-limiting symptoms during the DBPCFC at Week 12

An analysis of covariance (ANCOVA) model of change from baseline in log transformed MTD at Week 12 DBPCFC will be fit with terms for treatment group, age subgroup (12-17 years, 18-55 years), log-transformed total IgE at Screening Visit 1, region, and log-transformed MTD at baseline. Missing MTD at Week 12 values due to intercurrent events will be imputed by screening MTD, and other missing data will be multiply imputed under MAR assumption. If model assumptions of homoscedasticity and normality are not met, then the Wilcoxon rank-sum test may be performed.

- Change from baseline in peanut-specific IgE and IgG4 at Week 12, Week 16 and Week 52

Summary statistics, including geometric means and standard deviations, will be presented for peanut-specific IgE and IgG4 along with changes from baseline by time point and treatment group. The summary statistics will be provided for observed data. Change from baseline in log-

transformed levels of peanut specific-IgE and peanut specific-IgG4 at Week 12 and Week 16 will be analyzed using an ANCOVA model with terms for treatment group, age group, region, log-transformed total IgE at Screening Visit 1, and log-transformed baseline peanut specific-IgE or peanut specific-IgG4. Values below the LLOQ or above the ULOQ will be imputed as half the LLOQ or as the ULOQ, respectively.

- Change from baseline (screening) in mean wheal diameters by SPT to peanut at Week 16, Week 56 and Week 68

The mean is the average of wheal orthogonal diameter and wheal longest diameter minus the average of the non-missing negative control diameters. In most instances, the negative control diameters are 0mm.

Summary statistics will be presented by treatment group at each visit for:

- Undiluted peanut SPT
- Change from baseline in undiluted peanut SPT (mm)
- Percent change from baseline in undiluted peanut SPT
- Average across all dilutions
- Change from baseline across average of all dilutions (mm)
- Percent change from baseline across average of all dilutions

The summary statistics will be provided for observed data. Change from baseline in SPT mean wheal diameter at Week 16 will be analyzed using an ANCOVA model with terms for treatment group, age subgroup, region, log-transformed total IgE at screening, and baseline mean wheal diameter. This analysis will be performed for undiluted peanut SPT and the average across all dilutions (i.e. 1/100000, 1/10000, 1/1000, 1/100, 1/10, and undiluted). Note that the undiluted values are included in the average across all dilutions.

The analyses of PRO endpoints are described in Section 2.10.

2.7 Safety analyses

All safety endpoints (i.e. AEs, laboratory data, vital signs, and ECG) will be summarized by treatment for all participants on the safety set. In addition, subgroup analysis by age (12-17 years, 18-55 years) will be evaluated for most of the safety endpoints listed in this section. The safety analyses will be performed by the treatment groups for the three pooling periods as described in Section 2.1.1.1, unless otherwise specified.

2.7.1 Adverse events (AEs)

Unless otherwise specified, summaries will include treatment-emergent adverse events (TEAEs) only. Treatment emergent adverse events are defined as events started after the first dose of study treatment and within 16 weeks after the last study treatment, or events present prior to the first dose of study treatment but increased in severity based on preferred term within 16 weeks after the last study treatment. All events that the investigator classifies as reactions associated to the DBPCFC or SPT will not be included in reporting of treatment-emergent AEs and may be reported separately. The number and percentage (with exact binomial 95% CIs) of participants with TEAEs will be summarized in the following ways:

- by treatment, primary system organ class and preferred term.
- by treatment, primary system organ class, preferred term and maximum severity.
- by treatment, Standardized MedDRA Query (SMQ) and preferred term.

Separate summaries will be provided for study treatment-related adverse events, death, serious adverse events, other significant adverse events leading to discontinuation and TEAEs with potential cause of accidental exposure or ingestion. If a participant reported more than one AE with the same preferred term, the AE with the greatest severity will be presented. If a participant reported more than one AE within the same primary system organ class, the participant will be counted only once with the greatest severity at the system organ class level, where applicable.

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on TEAEs which are not SAEs with an incidence greater than a certain threshold based on the final database and on TESAEs and SAEs suspected to be related to study treatment will be provided by system organ class and PT on the safety set population.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

a. a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
b. more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non-SAE has to be checked in a block e.g., among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

2.7.1.1 Exposure-Adjusted Incidence Rate

The crude incidence rate is defined as the percentage of participants with a specific adverse event divided by the total number of participants in each study group.

Due to expected differences in exposure and follow-up due to varied duration of study participation between participants, adverse event incidence rates will be provided as “exposure adjusted AE incidence rates” in addition to the crude incidence.

The EAIR is defined as the number of participants with a specific event divided by the total exposure-time among the participants in the study group. That is, the EAIR is calculated as:

$EAIR = n / \Sigma\ t_i$, where n is the number of participants having the i^{th} type event, and t_i is a participant's exposure time and defined as the shortest of the following:

1. time to the first episode of the i^{th} type event (if the event occurs),
2. time to the end of observation period, or
3. duration of study treatment plus the 16-week washout period (approximately corresponding to five half-lives) after last treatment dose

The total exposure time of all participants in a treatment group is $\Sigma\, t_i$. The EAIR is interpreted as the number of events occurring in a population per unit time. The exact Poisson 95% confidence interval for the EAIR will be provided as well, where an exact 100*(1-α)% confidence interval will be derived as follows (Garwood 1936, Sahai and Khurshid 1993):

- Lower confidence limit $L = \frac{0.5C_{\alpha/2,2n}}{\sum t_i}$ for n>0, 0 otherwise,
- Upper confidence limit $U = \frac{0.5C_{1-\alpha/2,2n+2}}{\sum t_i}$,

where $C_{\alpha,k}$ is the αth quantile of the Chi-square distribution with k degrees of freedom.

Whenever applicable, exposure adjusted incidence rates will be provided for the type as below:

1. TEAE: Primary SOC level, PT level

2.7.1.2 Adverse events of special interest / grouping of AEs

AEs of special interest (AESI) are defined in the latest version of the compound electronic Case Retrieval Strategy (eCRS). The classification reflects the safety topics of interest identified in the current version of the QGE031 Development Safety Profiling Plan and may be updated based on review of accumulating data. The number and percentage of participants with treatment emergent AEs of special interest will be summarized by risk category, PT and treatment. The search criteria in the latest eCRS corresponding the MedDRA version at the database lock will be used and reported in the CSR:

- Hypersensitivity reactions (including anaphylaxis)
- Cardiovascular and Cerebrovascular (CCV) events
- Neoplastic conditions
- Injection site reactions
- Serum Sickness
- Eosinophilic Conditions / Churg-Strauss Syndrome
- Parasitic (Helminthic) infections
- Thrombocytopenia

Injection Site Reactions

For treatment emergent injection site reactions (ISR), besides the overall summary table for AESI, number of participants with recurrence of ISR (single ISR, 2-3 ISRs, > 3 ISRs) will be summarized by treatment.

Adjudicated AEs

From the AESIs listed above, the following AEs will be adjudicated by the independent committee. The adjudicated events will be listed, and a summary table may be provided following the adjudication.

- Anaphylaxis
- Cardiovascular and Cerebrovascular (CCV) events
- Neoplastic conditions

In addition, summaries will be provided by subgroups of age and allergy status (mono-sensitized and poly-sensitized).

2.7.2 Deaths

The number of deaths resulting from TEAEs will be summarized by SOC and PT. All the deaths in the clinical database including those occurring during screening will be listed.

2.7.3 Laboratory data

The descriptive summary statistics (mean (95% CI), median, first and third quartiles, minimum and maximum) for change from baseline to each study visit and maximum/minimum value will be presented. These descriptive summaries will be presented by test group, laboratory test and treatment group. Change from baseline will only be summarized for participants with both baseline and post baseline values for quantitative parameters, and the maximum/minimum value could come from post-baseline scheduled, unscheduled or premature discontinuation visits. For categorical parameters, frequencies by categories at each visit will be summarized.

The laboratory values below Lower Level of Quantification (LLOQ) or above Upper Level of Quantification (ULOQ) will be imputed as LLOQ or ULOQ in the summary tables, respectively. The numerical part of the reported result will be treated as the actual LLOQ or ULOQ. These laboratory values will be displayed in listings using the standard unit with the reported sign ("<" or ">").

For selected laboratory parameters, abnormalities occurring at any time-point from scheduled, unscheduled and premature discontinuation visits considering all post-baseline on-treatment data will be summarized. Where normal ranges are available, abnormalities in laboratory data may be listed by, participant, and visit/time.

The number of participants with worsening abnormality during the study will be summarized by treatment. A case is considered as worsening abnormality if the value at baseline and at least one post-baseline value during the study is worse than baseline.

2.7.4 Other safety data

2.7.4.1 ECG

Summary statistics (absolute values and change from baseline) for all ECG parameters will be provided by treatment and time point; Participants with notable abnormalities in ECG data will be listed by treatment group and visit/time. The following will be considered as notable values for adults: QT > 500 msec; QTcF > 450 msec (males), QTcF > 460 msec (females); PR > 250 msec. The following will be considered as notable ECG values for adolescents: QTcF >450 msec (males), QTcF >460 msec (females); PR >250 msec.

For ECG parameters, the number and percentage of patients with worsening abnormalities occurring post-baseline will be summarized using shift table by treatment group. The definition worsening is the same as laboratory values.

2.7.4.2 Vital signs

Summary statistics will be provided by vital sign, treatment group and visit/time as appropriate. Change from baseline will only be summarized for participants with both baseline and post-baseline values. Participants with notable vital signs as defined below will be listed.

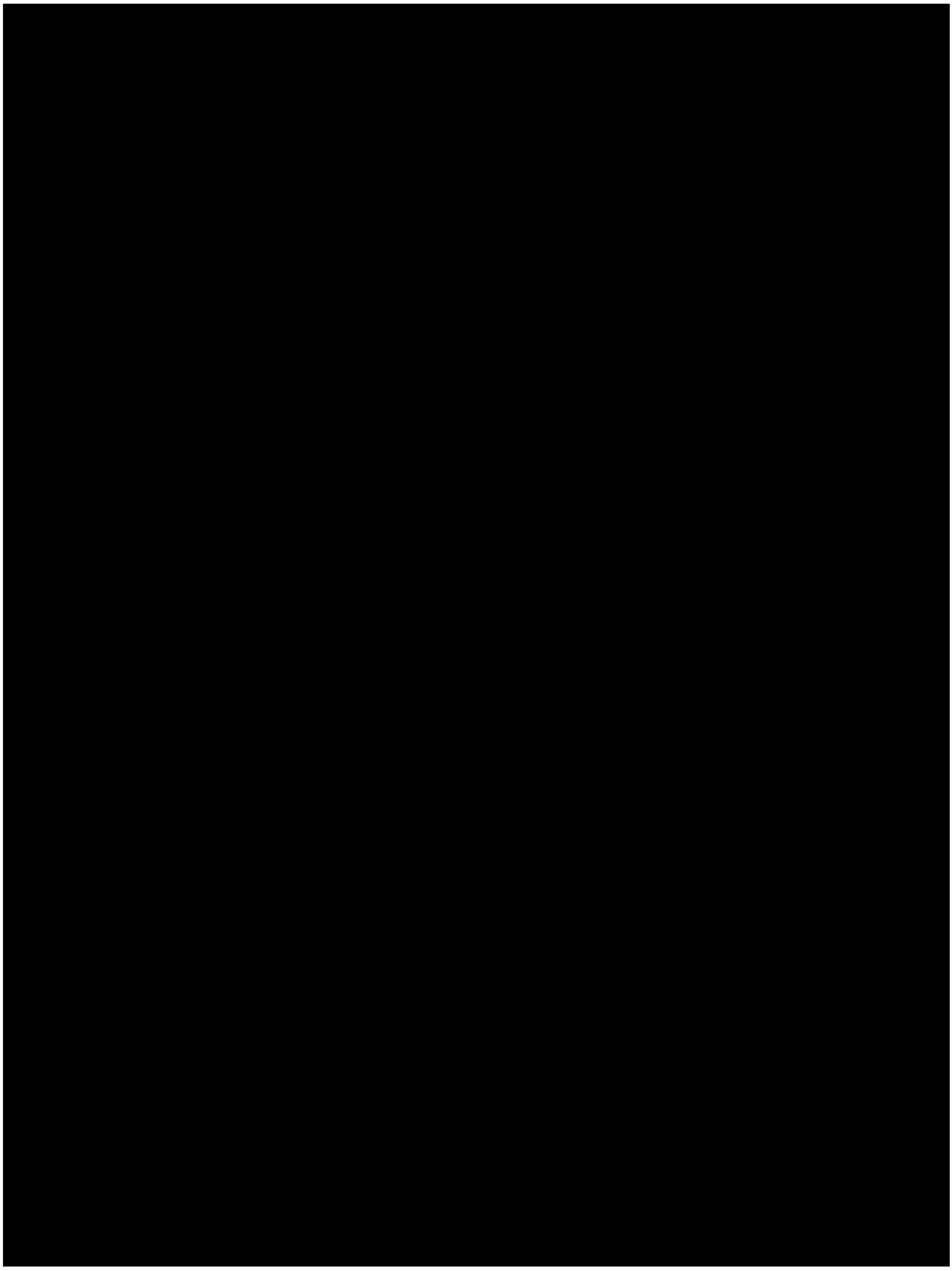
For adults:

- Hypertension (systolic blood pressure of ≥ 140 mmHg and/or diastolic blood pressure of ≥ 90 mmHg) or hypotension (systolic blood pressure of < 90 mmHg and/or diastolic blood pressure of < 60 mmHg)
- Pulse rate below 60 bpm (bradycardia) or above 100 bpm (tachycardia)

For adolescents:

See Section 5.3 for upper and lower limits for vital signs.

[redacted]



2.10 Patient-reported outcomes

2.10.1 Patient-reported outcome assessments

PRO data will be analyzed in accordance with the secondary endpoint described in Section 1.2. The following measures are being administered based on age group and responder type:

- FAQLQ-CF: Food Allergy Quality of Life Questionnaire – Child Form (completed by children aged 8-12)
- FAIM-CF: Food Allergy Independent Measure – Child Form (completed by children aged 8-12)
- FAQLQ-TF: Food Allergy Quality of Life Questionnaire – Teenager Form (completed by adolescents aged 13-17)
- FAIM-TF: Food Allergy Independent Measure – Teenager Form (completed by adolescents aged 13-17)
- FAQLQ-AF: Food Allergy Quality of Life Questionnaire – Adult Form (completed by adults aged 18-55)
- FAIM-AF: Food Allergy Independent Measure – Adult Form (completed by adults aged 18-55)
- FAQLQ-PF: Food Allergy Quality of Life Questionnaire – Parental Form (completed by same parent/caregiver of children aged 0-12)
- FAQL-PB: Food Allergy Quality of Life – Parental Burden Questionnaire (completed by same parent/caregiver of children aged 0-17)

- SF-36v2 Acute Version – Medical Outcomes Study 36-Item Short Form Version 2 Acute Version (recall period is past week) (completed by adults aged 18-55)

Table 2-1 Administration of questionnaires according to participant age.

Age Group / Respondent Type	Questionnaire	Day 1 Randomization	Week 12 (10 Days Before D1 OFC)	Week 12 (3 Days After D2 OFC)	Week 52 (10 Days Before D1 OFC)	Week 52 (3 Days After D2 OFC)
Children aged 12	FAQLQ-CF	x	x	x	x	x
	FAIM-CF	x	x	x	x	x
Teenagers aged 13-17	FAQLQ-TF	x	x	x	x	x
	FAIM-TF	x	x	x	x	x
Adults aged 18+	FAQLQ-AF	x	x	x	x	x
	FAIM-AF	x	x	x	x	x
	SF-36v2	x		x		x
Parents/ Caregivers of Children aged 13-17	FAQL-PB	x	x	x	x	x
Parents/ Caregivers of Children aged 12	FAQLQ-PF	x	x	x	x	x

The Food Allergy Quality of Life Questionnaire (FAQLQ) is a self-reported instrument intended to assess the effect of food allergy on the participant's health-related quality of life (HRQoL. The FAQLQ-CF (aged 8-12) (Flokstra – de Blok, DunnGalvin et al. 2009), FAQLQ-TF (aged 13-17) (Flokstra – de Blok, DunnGalvin et al. 2008) and FAQLQ-AF (≥18 years of age) (Flokstra-de Blok, van der Meulen et al. 2009) are self-administered, validated food allergy-specifc HRQoL questionnaires. The FAQLQ-PF is completed by parents of children aged 0 - 12 with food allergy (DunnGalvin, de BlokFlokstra et al. 2008).

Scoring of PRO Instruments

The number of items and domains varies by FAQLQ instrument administered. Each question is scored on a 7-point scale (coded as 1-7 in analysis, with a higher level indicating greater impairment in HRQoL). The total score is the arithmetic average of all completed items. Domain scores are calculated similarly for each form as following:

FAQLQ-CF: Emotional impact (EI) (item no. 19-24), Allergen avoidance (AA) (item no. 4, 6-10, 15), Risk of accidental exposure (RAE) (item no. 11, 13, 14, 16, 17), Dietary restriction (DR) (item no. 1-3, 5, 12, 18)

FAQLQ-TF: Emotional impact (EI) (item no. 5, 12, 19-23), Allergen avoidance and dietary restrictions (AADR) (item no. 1-4, 6-10, 16), Risk of accidental exposure (RAE) (item no. 11, 13-15, 17, 18)

FAQLQ-AF: Emotional impact (EI) (item no. 5, 24-29), Allergen avoidance and dietary restrictions (AADR) (item no. 1-4, 6, 8-12, 20), Risk of Accidental Exposure (RAE) (item no. 7, 13-18, 21), Food allergy related health (FAH) (item no. 19, 22, 23)

FAQLQ-PF: Emotional impact (EI) (item no. 2, 6, 7, 9-11, 23-28, 30), Food anxiety (FA) (item no. 1, 4, 5, 16, 17, 20, 21, 29), Social and dietary limitations (SDL) (item no. 3, 8, 12-15, 18, 19, 22).

If more than one item in any domain is missing, a domain score should not be calculated for that case. A total score can still be calculated if 20% or fewer of the items are missing.

The FAIM reflects the participant's perceived food allergy severity and food allergy-related risk (van der Velde, Flokstra-de Blok et al. 2010). Each question is scored on a 7-point scale (coded as 1-7 in analysis, with a greater score indicating a higher level of perceived risk or chance of adverse events occurring). The total score is the arithmetic average of all completed items. If less than 80% of the items within the score are complete, it will not be calculated. Questions 4 and 5 (effectively managing a reaction, receiving sufficient help from others) must be reverse coded.

A threshold of 0.45 points has been suggested and used as a minimal important difference (MID) to interpret clinical relevance (Dunn Galvin, Cullinane et al. 2010). For FAQLQ and FAIM forms, the 0.45 threshold will be used to interpret the clinical relevance of group-level differences and changes, in addition to within-patient changes.

The FAQL-PB is a self-administered, disease-specific instrument developed to measure the effect of pediatric food allergy on HRQoL for caregivers (Cohen, Noone et al. 2004). The instrument includes 17 items (assessing social, dietary and emotional impacts) on the impact of having a child with food allergy on the parents themselves using a 7-point Likert scale ranging from 0 (not limited/troubled) to 6 (extremely limited/troubled). The total score is the arithmetic average of all completed items.

The SF-36v2® Health Survey is a 36-item instrument that measures generic health-related quality of life. It is designed for use in surveys of general and specific populations, health policy evaluations and clinical practice and research. Two forms of this instrument are available and this study will use the one-week recall (acute) form. The SF-36v2 contains 8 scales and 2 component summary indices evaluating physical, social and emotional functioning in addition to general health perceptions and mental health.. Responses to items allow for direct calculation of scale scores, while the physical component summary (PCS) and mental component summary (MCS) scores are computed from weighted scale scores. The SF-36v2 scale and composite scores can be converted to a T-score metric, allowing for norm-based scores derived from responses to a 2009 survey conducted by QualityMetric.[4] For all scales and summary measures, higher scores indicate better health outcomes.

The SF-6Dv2 will be derived from the SF-36v2 for health economic evaluations (Brazier, Mulhern et al. 2020). The SF-6Dv2 captures the impact of diseases and conditions on social activities and depression/nervousness. The SF-6Dv2 questionnaire uses a standardized health state descriptive system consisting of 6 dimensions, including social limitations, which was deemed essential to cover in food allergy. All dimensions are expressed over 5 levels, except for pain, which uses 6 levels, allowing for the description of 18,750 different health states. Health states are described using a combination of 6 digits each expressing the level for a dimension. For example, state "111111" indicates perfect health and state "555655" indicates the worst possible health status. The resulting SF-6Dv2 index, scored from 0.0 (worst health

state) to 1.0 (best health state), can be used in the assessment of the quality adjusted life years (QALYs) and the cost-effectiveness of various health care interventions.

2.10.1.1 Patient-reported outcome analysis

All PRO scores will be summarized descriptively at Day 1, before and after the DBPCFC at Week 12 and Week 52 by treatment. Tabular summaries will present continuous PRO total and domain scores, plus change from baseline scores, as per Section 2.11.1. Summaries for change from baseline in FAQLQ and FAIM domain and total scores at Week 12 (before and after the DBPCFC) will also be provided by response status and treatment during DBPCFC at Week 12 (Responder/Non-responder at 600 mg of peanut protein). Change from baseline in SF-6Dv2 social function and mental health dimensions at Week 12 will be summarized by response status and treatment.

Change from baseline after Week 12 OFC Day 2 in FAQLQ and FAIM total scores will be analyzed using ANCOVA models with treatment group and region as fixed effects, and log-transformed total IgE at screening as well as respective PRO baseline value as covariates. Least Squares Mean differences, corresponding 95% CIs and p-values will be presented. All p-values presented from PRO inferential analyses are unadjusted and should be used to aid interpretation of the data only. Change from baseline before Week 12 OFC Day 1 in FAQLQ and FAIM total scores will be analyzed in a similar fashion.

All analyses will be performed on the FAS using observed data with no imputation of missing data.

2.11 Biomarkers

Peanut-specific IgE and IgG4 and SPT are analyzed as other secondary endpoints and described in Section 2.6.2. [redacted]

Additional details will be provided in a separate analysis plan targeting the analyses of biomarkers.

[redacted]

2.13 Interim analysis

There will be two analyses before the final database lock, in addition to the DMC analyses as described below.

Interim Analysis:

At the start of the study, recruitment will be restricted to 12-55 year old participants. When approximately 60 adolescent participants (defined as 12 -17 years of age) have completed all week 12 assessments, an interim analysis on [redacted] safety [redacted] data [redacted] will be performed (Section 2.9). Safety data (semi-blinded or unblinded) will only be generated for DMC review. [redacted]

Primary Analysis at Week 12:

The Primary Analysis will be performed once all participants have reached Week 12 in the study and completed its assessments or prematurely withdrawn from the study prior to Week 12. Formal testing of primary and key secondary endpoints will be performed according to testing strategy specified in Figure 2-1 with full alpha. Since the primary and key secondary objectives will be performed only for the Primary Analysis, adjustment to the overall type I error rate is not required. In addition, a Dose-Exposure-Response model to describe the responder rate and the factors impacting the response will be developed.

Table 2-2 provides the list of results that may be presented for the Primary Analysis at Week 12.

Table 2-2 Results presented at the Primary Analysis

Category	Analysis Description
OFC	Effect of 12 weeks of Ligelizumab: Analysis of the Proportion of Participants who Tolerated a Single Dose of >= 600 mg of Peanut Protein without Dose-Limiting Symptoms during DBPCFC at Week 12 using Logistic Regression Model (Full Analysis Set)
OFC	Effect of 12 weeks of Ligelizumab: Analysis of the Proportion of Participants who Tolerated a Single Dose of >= 1000 mg of Peanut Protein without Dose-Limiting Symptoms during DBPCFC at Week 12 using Logistic Regression Model (Full Analysis Set)
OFC	Effect of 12 weeks of Ligelizumab: Analysis of the Proportion of Participants who Tolerated a Single Dose of 3000 mg of Peanut Protein without Dose-Limiting Symptoms during DBPCFC at Week 12 using Logistic Regression Model (Full Analysis Set)
OFC	Effect of 4 weeks of Ligelizumab: Analysis of the Proportion of Participants who Tolerated a Single Dose of >= 1000 mg of Peanut Protein without Dose-Limiting Symptoms during DBPCFC at Week 12 (Full Analysis Set)
OFC	Summary of Proportion of Participants who Tolerated at Least the Single Dose of >=600 mg Peanut Protein without Dose-Limiting Symptoms during DBPCFC at Week 12 (Full Analysis Set)
OFC	Proportion of Participants who Tolerated a Single Dose of >= 600 mg of Peanut Protein without Dose-Limiting Symptoms during DBPCFC at Week 12 by Timing of Study Termination Notice (Full Analysis Set)
OFC	Marginal Proportion of Participants who Tolerated a Single Dose of >= 600 mg of Peanut Protein without Dose-Limiting Symptoms during DBPCFC at Week 12 (Full Analysis Set)
OFC	Odds of Participants Developing Less Severe Level of Symptoms Evaluated by Maximum Severity of Symptoms at Any Challenge Dose of Peanut Protein up to and including 1000 mg during the DBPCFC Conducted at Week 12 (Full Analysis Set)

SPT	Summary of Skin Prick Test (undiluted peanut protein) and change from baseline by treatment group and across visits
SPT	Summary of Skin Prick Test (average dilutions peanut protein) and change from baseline by treatment group and across visits
PRO	FAQLQ-PF: Food Allergy Quality of Life Questionnaire – Parental Form (completed by parents of children aged 12)
PRO	FAQLQ-TF: Food Allergy Quality of Life Questionnaire – Teenager Form (completed by adolescents aged 13-17)
PRO	FAIM-TF: Food Allergy Independent Measure – Teenage Form (completed by adolescents aged 13-17))
PRO	FAQLQ-AF: Food Allergy Quality of Life Questionnaire – Adult Form (completed by adults aged 18-55)
PRO	FAIM-AF: Food Allergy Independent Measure – Adult Form (completed by adults aged 18-55)
Adverse Event	Number of adverse events due to Accidental Exposure (plus respective listing)
Event	Use of Epinephrine During and Outside of OFC
Safety	Overview of Treatment Emergent Adverse Events (Safety Set)
Safety	Exposure-adjusted TEAE tables for the entire population (Safety Set)
Listing	Deaths and SAEs by system organ class and preferred term (plus respective listing)
Safety	Treatment-emergent SAEs by primary system organ class
Safety	Treatment-emergent SAEs by preferred term (plus respective listing)
Safety	TEAEs by primary system organ class
Safety	TEAEs, by preferred term (plus respective listing)
Safety	TEAEs by preferred term and maximum severity
Safety	TEAEs leading to discontinuation by primary system organ class
Safety	TEAEs leading to discontinuation by preferred term (plus respective listing)
Safety	Treatment-emergent Adverse Events of Special Interest (AESI), by standardized MedDRA term (plus respective listing)
Safety	Number of AEs (n, %)
Safety	Newly occurring or worsening chemistry and hematology abnormalities (plus respective listing)
Safety	Newly occurring or worsening vital sign abnormalities (plus respective listing)
Safety	Adverse events for screen failures
Safety	Number of Subjects with Adjudicated Anaphylaxis Events (Safety Set) (plus respective listing)
Safety	Number of Subjects with Adjudicated CCV Events (Safety Set) (plus respective listing)
Safety	Number of Subjects with Adjudicated Neoplastic Events (Safety Set) (plus respective listing)
Safety	Number of AEs requiring Epinephrine treatment (n, %)
Disposition	Participant Disposition at Treatment Period (Randomized Set)
Set	Analysis Sets (Randomized Set)
Baseline	Subject Demographics (Randomized Set)
Baseline	Subject Baseline Characteristics (Randomized Set)

In addition, the following analyses will be presented.

DMC analyses:

The DMC will assess, at defined intervals, the progress of a clinical trial, safety data, and critical efficacy variables and recommend to the sponsor whether to continue, modify, or terminate a trial, as defined in the DMC charter. Analysis of selected efficacy variables will be provided only for the purpose of helping for safety evaluation. Such assessments do not inflate the Type I error for the primary efficacy hypothesis testing and thus no adjustment for multiplicity is required.

3 Sample size calculation

A total sample size of approximately 486 randomized participants is originally planned to achieve sufficient power for the primary and key secondary endpoints, and it provides adequate precision in estimating AE rates in this study. Due to the halt of recruitment, the sample size is now 211. This reduced sample size provides sufficient power for the primary endpoint.

Participants will receive one of three treatments (ligelizumab 240mg SCq4w, ligelizumab 120mg SCq4w, or placebo SCq4w) and will be randomized into five treatment arms with a ratio 2:2:2:2:1. At the start of the study, recruitment was restricted to participants 12-55 years old. Due to the recruitment halt, no other age groups were recruited.

3.1 Primary endpoint(s)

The primary objective is to demonstrate that ligelizumab doses are superior to placebo in responder rate at a level of 600 mg peanut protein (1044 mg cumulative tolerated dose) after 12 weeks of treatment.

The responder rate at 600 mg dose of peanut protein in placebo and ligelizumab groups have been assumed approximately 20% and at least 80%, respectively, for participants with evaluable DBPCFC at Week 12. These assumptions are considered highly clinically relevant based on the PALISADE trial conducted in similar populations (Vickery et al 2018). In the PALISADE trial, the responder rate at 600 mg single challenge dose of peanut protein in placebo and AR101 were 4.3% (95% CI: 1.9%, 9.7%) and 84.5% (95% CI: 79.9%, 88.1%), respectively. Approximately 15% of participants were assumed to discontinue treatment before week 12 and will be considered non-responders in power calculations. Therefore, the responder rate at 600 mg single dose of peanut protein in placebo and ligelizumab were assumed 17% and 68%, respectively, for randomized participants. For the purpose of evaluating the power for the primary trial objectives, it was assumed that none of the secondary null hypotheses would be rejected and the full alpha level of 0.025 (one-sided) would be equally split across the primary

hypotheses for each dose (i.e., 0.0125 one-sided for each testing). Under the outlined assumptions, a two group Fisher's exact test with a 0.0125 one-sided significance level will have above 99% power to detect the difference of responder rate between each ligelizumab group and placebo, when the sample size is 108 participants on each ligelizumab dose and 54 participants on placebo. Due to halt of recruitment, the sample size is expected to be around 211 (around 46 participants on each ligelizumab dose and 23 participants on placebo). This reduced sample size provides 96% power to detect the difference of responder rate between each ligelizumab group and placebo at a 0.0125 one-sided significance level. These power calculations are an approximation of the power achieved with the logistic regression approach. The table below shows the sensitivity of the power to deviations from the assumptions.

Table 3-1 Sensitivity of the power for each dose for the primary variable

RR in each QGE031 group with evaluable DBPCFC at Week 12	**RR in placebo group with evaluable DBPCFC** at Week 12	**Drop-out rate**	**RR in each QGE031 group**	**RR in placebo group**	**Sample size in each QGE031 group**	**Sample size in placebo**	**Power**
80%	20%	15%	68%	17%	108	54	>99.9%
80%	20%	15%	68%	17%	46	23	96%
70%	20%	15%	59.5%	17%	108	54	99.8%
70%	20%	15%	59.5%	17%	46	23	86%

RR = responder rate. Approximately 15% of participants were assumed to discontinue treatment before week 12 and be considered as non-responders in power calculations. Power results were calculated with PASS version 11.

3.2 Secondary endpoint(s)

If statistical significance is achieved in the primary test, the tests for the key secondary variables included in the testing strategy will be performed. The local significance level for each key secondary null hypothesis will be determined based on the closed testing procedure shown in Figure 2-1.

Assuming a treatment discontinuation rate of 15% before week 12 and responder rates in each treatment group as shown in the table below, the local power (unconditional) of each hypothesis was estimated using 10,000 simulations with package gMCP in R 3.6.1. The assumption for maximum severity of symptoms is based on the PALISADE trial conducted in similar populations (Vickery et al 2018). The hypothesized response to 4 weeks of ligelizumab is supported by results reported in Savage et al 2012 [redacted]. According to Savage et al 2012, omalizumab increased the median tolerated threshold dose of peanut protein from 80 mg at baseline to 6500 mg at Week 5 approximately. The underlying distribution of difference in responder rate between ligelizumab and placebo is assumed to be multivariate normal in power simulations. To be conservative, correlation between endpoints was ignored for power simulation. The originally proposed sample size of 486 ensured sufficient power for primary and key secondary endpoints.

Table 3-2 Power of each dose for the analyses of primary and key secondary variables in the base case scenario

	Assumptions of treatment effect				Local power (unconditional)	
Variable		**QGE031 240 mg**	**QGE031 120 mg**	**Pbo**	**QGE031 240 mg**	**QGE031 120 mg**
RR of tolerating 600 mg peanut protein	RR* (RR in rand)	80% (68%)	80% (68%)	20% (17%)	>99.9%	>99.9%
RR of tolerating 1000 mg peanut protein	RR* (RR in rand)	70% (59.5%)	70% (59.5%)	20% (17%)	>99.9%	>99.9%
RR of tolerating 3000 mg peanut protein	RR* (RR in rand)	60% (51%)	60% (51%)	15% (12.75%)	>99.9%	>99.9%
Maximum severity of symptoms at any challenge dose up to and including 1000 mg#	% None	38%	38%	2%	>99.9%	>99.9%
	% Mild	32%	32%	28%		
	% Moderate	25%	25%	59%		
	% Severe	5%	5%	11%		
RR of tolerating 1000 mg peanut protein with 4 weeks of ligelizumab treatment	RR^	50%	45%	20%	98.3%	93.0%

RR=responder rate; Pbo=placebo; The proposed number of participants for comparisons between ligelizumab and placebo are 108 in each ligelizumab arm and 54 in placebo arm.
*RR is based on the anticipated proportion of responders in participants with evaluable DBPCFC at Week 12. Approximately 15% of participants are assumed to discontinue treatment before week 12 and considered non-responders for randomized population in power calculation (RR in rand). #For maximum severity of symptoms, the assumed proportion of participants in each category is based on randomized population. Power calculations were performed in R 3.6.1 with package gMCP.
^RR is based on the anticipated proportion of responders in participants with evaluable DBPCFC at Week 12. Same treatment effect is expected for randomized population in power calculation, assuming participants had taken study treatment at Week 8 and conducted DBPCFC at Week 12.

3.2.1 Precision for adverse events

Table 3-3 Precision levels (95% CI) for sensitivity of safety incidence rate for overall population

Event rate (proportion of participants with an event)	Number of participants	95% CI
0%	486	(0%, 0.8%)
1%	486	(0.3%, 2.3%)
2%	486	(1%, 3.7%)
0%	211	(0%, 1.7%)
1%	211	(0.1%, 3.5%)
2%	211	(0.6%, 4.9%)

4 Change to protocol specified analyses

On 31 March 2023, Novartis made a strategic decision to close recruitment in study CQGE031G12301 and terminate the study early. As a result, there are several changes to the specified analyses in the protocol, namely:

- Instead of 486 participants, the sample size is 211.

- There will be no recruitment of children under age 12, and the age-groups of interest are now 12-17 and 18-55.
- For participants who have reached Screening Visit 2 Part 1 and not completed Week 52 visit at the time of halt recruitment will need to complete Screening Visit 2 Part 2 to determine eligibility. If they are eligible, they will either continue treatment and perform assessments until Week 12 Part 2 or will stop treatment at the next scheduled visit if they have already completed Week 12 Part 2. This leads to limited number of participants reaching Week 52. Hence, non-inferiority test for sustainability of treatment effect at Week 52 compared to Week 12 is removed.
- Week 52 interim analysis will not be conducted.

There are also some changes made to simplify the analyses:

- Since the number of COVID-19 related intercurrent events is expected to be minimal, references to COVID-19 related intercurrent events were removed.
- Handling of intercurrent events for the secondary estimands were changed to be consistent with the primary estimand strategy.
- The two-dimensional tipping point analysis for the primary estimand was removed.

In analysis for primary estimand and secondary estimands(1, 2 and 4), log-transformed MTD at screening DBPCFC was added as covariate in logistic regression models.

The protocol mentions the summary measure of supplementary estimand is relative risk generated with a marginal standardization method. In SAP, relative risk is changed to risk difference to facilitate clinical interpretation.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

No imputation of missing/partial start or end study drug date. If missing, the time of study end date will be imputed to 00:00:00.

5.1.2 AE date imputation

Rules for imputing the AE end date:

- If the AE end date month is missing, then the imputed end date should be set to the earliest of the study end date, 31DECYYYY or date of death.
- If the AE end date day is missing, then the imputed end date should be set to the earliest of the study end date, last day of the month or date of death.
- If AE year is missing or AE is ongoing, then the end date will not be imputed.

Rules for imputing the AE start date:

1. If imputing end dates, then this should be done prior to calculating imputed start dates.

Before imputing AE start date, find the AE start reference date.

- If the (imputed) AE end date is complete and the (imputed) AE end date < treatment start date, then AE start reference date = min(informed consent date, earliest visit date).
- Otherwise, AE start reference date = treatment start date.

Impute AE start date:

- If the AE start date year value is missing, then the date uncertainty is too high to impute a rational date. Therefore, if the AE year value is missing, then the imputed AE start date is set to NULL.
- If the AE start date year value is less than the treatment start date year value, then the AE started before treatment. Therefore:
  - If AE month is missing, then the imputed AE start date is set to the mid-year point (01JulYYYY).
  - Otherwise, if AE month is not missing, then the imputed AE start date is set to the mid-month point (15MONYYYY).
- If the AE start date year value is greater than the treatment start date year value, then the AE started after treatment. Therefore:
  - If the AE month is missing, then the imputed AE start date is set to the year start point (01JanYYYY).
  - Otherwise, if the AE month is not missing, then the imputed AE start date is set to the later of month start point (01MONYYYY) or AE start reference date + 1 day.
- If the AE start date year value is equal to the treatment start date year value:

- If the AE month is missing, then the imputed AE start date is set to the AE reference start date + 1 day.
- If the AE month is less than the treatment start month, then the imputed AE start date is set to the mid-month point (15MONYYYY).
- Otherwise, if the AE month is equal to the treatment start date month or greater than the treatment start date month, then the imputed AE start date is set to the later of (month start point (01MONYYYY), AE start reference date + 1 day).

If complete (imputed) AE end date is available and the imputed AE start date is greater than the (imputed) AE end date, then imputed AE start date should be set to the (imputed) AE end date.

5.1.3 Concomitant medication date imputation

Rules for imputing the CM end date (including on-going records):

a. If imputing end dates, this should be done prior to calculating imputed start dates.
b. When the medication is ongoing at the end of the study, no numeric end date is derived.
c. If the end date is completely missing no numeric end date is derived.
1. If CM end day is missing and CM month/year are non-missing, then impute CM date as the minimum of study end date and the last day of the month.
2. If CM end day/month are missing and CM year is non-missing, then impute CM date as the minimum of study end date and the end of the year (31DECYYYY).
3. If imputed CM end date is less than the complete CM start date, use the complete CM start date as the imputed CM end date.

Rules for imputing the CM start date:

1. If imputing end dates, then this should be done prior to calculating imputed start dates.

- If the CM start date year value is missing, then the imputed CM start date is set to one day prior to *Treatment start date (TR01SDT)*.
- If the CM start date year value is less than the *Treatment start date (TR01SDT)* year value, then the CM started before treatment. Therefore;
  - If the CM month is missing, then the imputed CM start date is set to the mid-year point (01JulYYYY).
  - Else if the CM month is not missing, then the imputed CM start date is set to the mid-month point (15MONYYYY).
- If the CM start date year value is greater than the *Treatment start date (TR01SDT)* year value, the CM started after treatment. Therefore;
  - If the CM month is missing, then the imputed CM start date is set to the year start point (01JanYYYY).
  - Else if the CM month is not missing, then the imputed CM start date is set to the month start point (01MONYYYY).
- If the CM start date year value is equal to the *Treatment start date (TR01SDT)* year value;
  - And the CM month is missing or the CM month is equal to the *Treatment start date (TR01SDT)* month, then the imputed CM start date is set to one day prior *Treatment start date (TR01SDT).*

- Else if the CM month is less than the *Treatment start date (TR01SDT)* month, then the imputed CM start date is set to the mid-month point (15MONYYYY).
- Else if the CM month is greater than the *Treatment start date (TR01SDT)* month, then the imputed CM start date is set to the month start point (01MONYYYY).

If complete (imputed) CM end date is available and the imputed CM start date is greater than the complete (imputed) CM end date, then imputed CM start date should be set to the complete (imputed) CM end date.

If there is no end date and ongoing check is not ticked, the CM will be considered as ongoing and included in the summary table.

5.2 AEs coding/grading

The MedDRA version which will be available at the time of database lock, will be used for the coding purpose of the adverse events.

5.3 Laboratory parameters derivations

Refer to Section 16.2 of the protocol for clinically notable laboratory values for hepatotoxicity.

Refer to Section 16.3 of the protocol for clinically notable laboratory values for nephrotoxicity

The following other specific criteria have been identified for this study:

- Platelets < 75 000/µL
- Any participant who has platelets < 75 000/µL after being randomized should discontinue study treatment.

For all other laboratory assessments, the Central Laboratory will flag laboratory values falling outside of the normal ranges on the Central Laboratory Report (which the investigator should review and sign-off) and the investigator will report any values considered clinically significant in the CRF.

Vital signs

Notable values for vital signs for adults are:

- heart rate of < 60 and > 100 bpm
- systolic blood pressure of < 90 and ≥ 140 mmHg
- diastolic blood pressure of < 60 and ≥ 90 mmHg

For participants aged 12 and older, the notable values are described in Table 5-1:

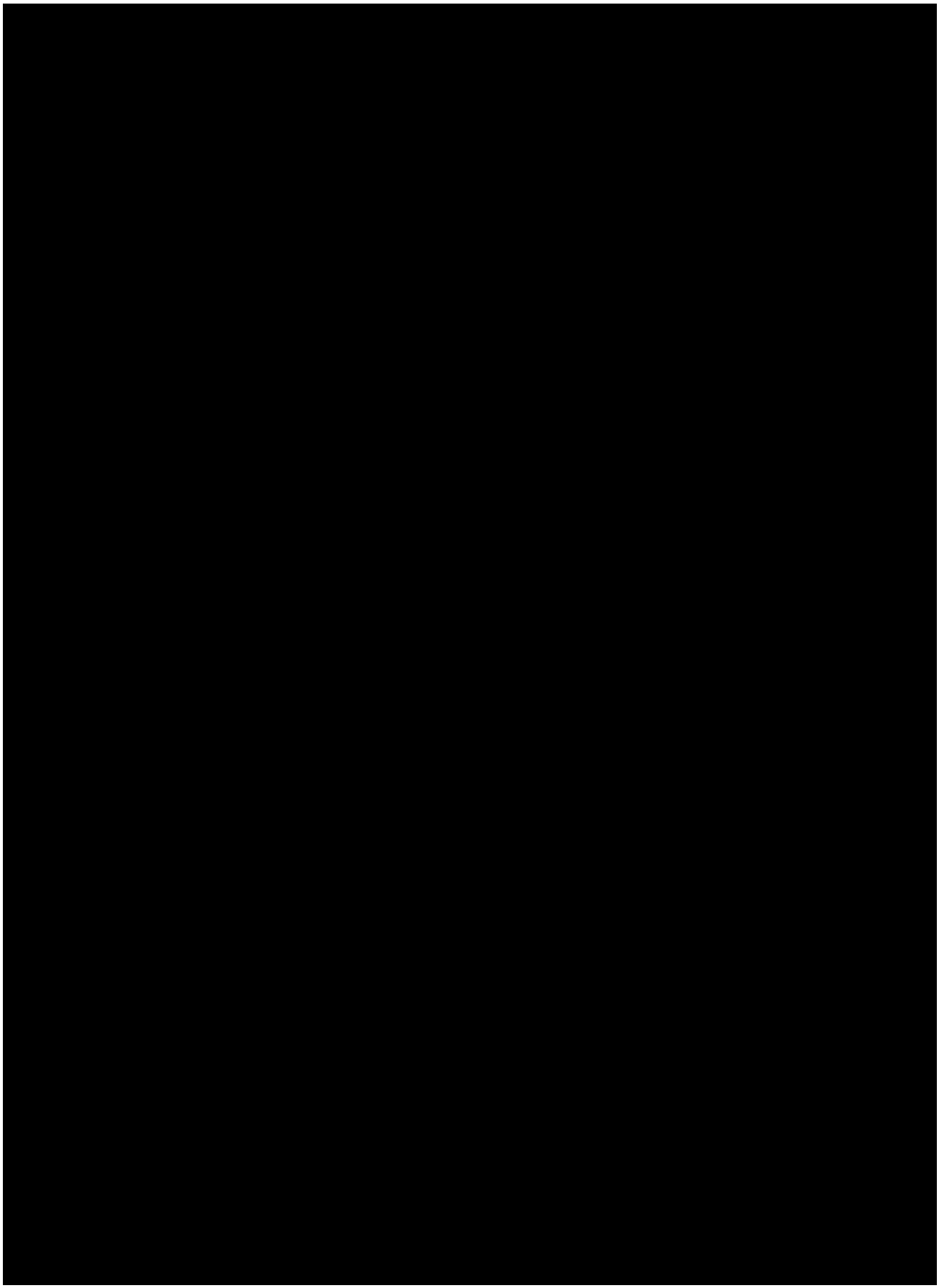
Table 5-1 Notable values requiring further evaluation for Heart Rate (HR) and Blood Pressure in children and adolescents

Age (years)	HR (bpm)		Blood Pressure (mmHg)	
	Low	High	Systolic	Diastolic
12	< 67	> 103	>= 114 for girls (>= 113 for boys)	>= 75
13 or 14	< 62	> 96	>= 120	>= 80
15 to 17	< 58	> 92	>= 120	>= 80

Heart Rate (HR) Adapted from Fleming et al 2011; Blood Pressure (BP) adapted from Flynn 2017

[redacted]

[redacted]



5.5 Rule of exclusion criteria of analysis sets

Table 5-2 Criteria leading to exclusion

Analysis Set	Criteria that cause subjects to be excluded
RAN	Not randomized
FAS	Not in RAN; Mistakenly randomized and no double-blind study drug taken
SAF	No double-blind study drug taken

6 Reference

Brazier, John & Mulhern, Brendan & Bjorner, Jakob & Gandek, Barbara & Rowen, Donna & Alonso, Jordi & Vilagut, Gemma & Ware, John. (2020). Developing a New Version of the SF-6D Health State Classification System From the SF-36v2: SF-6Dv2. Medical Care. 58. 10.1097/MLR.0000000000001325.

Bretz F, Maurer W, Brannath W, et al (2009) A graphical approach to sequentially rejective multiple test procedures. Statistics in Medicine; 28:586-604.

Cohen BL, Noone S, Muñoz-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 Nov;114(5):1159-63. doi: 10.1016/j.jaci.2004.08.007. PMID: 15536425.

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 Jun;38(6):977-86. doi: 10.1111/j.1365-2222.2008.02978.x. Epub 2008 Apr 23. PMID: 18435800.

Fleming S, Thompson M, Stevens R, et al (2011) Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. Lancet; 377: 1011-8.

Flokstra-de Blok BM, Dubois AE, Vlieg-Boerstra BJ, Oude Elberink JN, Raat H, DunnGalvin A, Hourihane JO, Duiverman EJ. Health-related quality of life of food allergic patients: comparison with the general population and other diseases. Allergy. 2010 Feb;65(2):238-44. doi: 10.1111/j.1398-9995.2009.02121.x. Epub 2009 Oct 1. PMID: 19796214.

Flokstra-de Blok BM, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JN, Duiverman EJ, Hourihane JO, Dubois AE. Development and validation of the self-administered Food Allergy Quality of Life Questionnaire for adolescents. J Allergy Clin Immunol. 2008 Jul;122(1):139-44, 144.e1-2. doi: 10.1016/j.jaci.2008.05.008. PMID: 18602570.

Flokstra-de Blok BM, van der Meulen GN, DunnGalvin A, Vlieg-Boerstra BJ, Oude Elberink JN, Duiverman EJ, Hourihane JO, Dubois AE. Development and validation of the Food Allergy Quality of Life Questionnaire - Adult Form. Allergy. 2009 Aug;64(8):1209-17. doi: 10.1111/j.1398-9995.2009.01968.x. Epub 2009 Feb 11. PMID: 19210345.

Flynn JT, Kaelber DC, Baker-Smith CM, et al (2017) Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents. Pediatrics; 140(3):e20171904.

Ge, M., Durham, L.K., Meyer, R.D., Xie, W. and Thomas, N., 2011. Covariate-adjusted difference in proportions from clinical trials using logistic regression and weighted risk differences. Drug information journal: DIJ/Drug Information Association, 45(4), pp.481-493.

ICH E9(R1) Harmonized Guideline: addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Final version on 20 November 2019.

Kawai N, Chuang-Stein C, Komiyama O & Li Y(2008) An Approach to Rationalize Partitioning Sample Size into Individual Regions in a Multiregional Trial. Drug Information Journal 42;(2) 139–47.

Savage JH, Courneya JP, Sterba PM, et al (2012) Kinetics of mast cell, basophil, and oral food challenge responses in omalizumab-treated adults with peanut allergy. J. Allergy Clin. Immunol. 130(5):1123-1129

van der Velde JL, Flokstra-de Blok BM, Vlieg-Boerstra BJ, Oude Elberink JN, DunnGalvin A, Hourihane JO, Duiverman EJ, Dubois AE. Development, validity and reliability of the food allergy independent measure (FAIM). Allergy. 2010 May;65(5):630-5. doi: 10.1111/j.1398-9995.2009.02216.x. Epub 2009 Oct 21. PMID: 19845570.

Vickery BP, Vereda A, et al (2018) AR101 Oral Immunotherapy for Peanut Allergy. N. Engl. J. Med.; 379(21):1991-2001