NOVARTIS

Clinical Development

CIGE025/Omalizumab/Xolair®

CIGE025E2305 / NCT03328897

A multicenter, randomized, double-blind, placebo-controlled phase III study to evaluate the efficacy and safety of Xolair® (omalizumab) in Chinese patients with chronic spontaneous urticaria who remain symptomatic despite antihistamine treatment

Statistical Analysis Plan (SAP)

Author: [redacted]

Document type: SAP Documentation

Document status: Final version 3.0 Amendment 2

Release date: 11-Nov-2019

Number of pages: 49

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
18-Jan-2017	Prior to FPFV	Creation of final version 1.0	N/A – First version	N/A
01-Sep-2017	Prior to DB lock	Final version 2.0	Systemic treatments list added Typos correction	All
28-Sep-2018	Prior to DB lock	Final version 2.1	Updates on analysis sets definition	Section 2.2
			Adding of notable value for platelets count	Section 2.8.3
			Adding of Protocol Deviations INCL03A and EXCL13 for FAS exclusion	Section 5.6
			Added of In-clinic at baseline	Section 2.3.2
			Precision on censoring of time to response in Time to ISS7 MID response	Section 2.7.2
			[redacted]	[redacted]
			Updated 'Change to protocol specified analyses' section	Section 4
			Update on DLQI for study week to analysis	Section 5.1
21-Mar-2019	Prior to DB lock	Final version 2.2	Adding of population to use: FAS	Section 2.14
			Population updated by SAF	Section 2.10
			Precision added for Prior urticaria therapies imputation	Section 5.2.3
			Precision added for latent TB status at baseline	Section 5.8
			[redacted]	[redacted]
			Statistical methodology details added	Section 5.5.2
			Per-protocol set population added	Section 2.2 / 5.6 Section 2.8.1.2

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			Precision added for China requirement	
11-Nov-2019	Prior to DB lock	Final version 3.0 Amendment 2	Duration of exposure updated for subjects who died. Baseline definition updated for subjects who did not take any study medication. On treatment period definition removed, since the definition is no longer required for the TFL shells.	Section 2.1.1 and 2.4.1
			PPS definition updated in order to exclude only Protocol Deviations with an impact on the primary endpoint. Also, 3 injections corrected to 3 doses of study drug. SAF definition updated in order to include all subjects with one dose of study medication taken. Also, corrected actual treatment derivation as per PPS derivation, i.e. 300 mg is two 150 mg injections at once (instead of one 300 mg injection).	Section 2.2
			Concomitant medications will be summarized by ATC class and preferred term. Prior medications not for CSU will be summarized by ATC class, preferred term and treatment group.	Section 2.4.2
			Baseline ISS7 definition updated in order to use the first treatment intake date, rather than the randomization date since both dates can be different. Added clarification that week 12 is derived from the 7 days prior the planned (not actual) day of the Week 12 visit (study days 78 - 84).	Section 2.5.1
			Rules for the primary endpoint updated in order to exclude all eDiary data collected on and after last dose +28 days for patients with less than three doses of study drug. Clarified weekly score derivations apply to all weeks (not just the 7 days prior to a study visit).	Section 2.5.3
			Removed ‘Analysis of the key secondary objective’ section, since there are no ‘key’ secondary objectives and added secondary endpoints.	Section 2.6.1

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			Clarified time to ISS7 MID is 'by Week 12' instead of 'during randomized treatment epoch'. Also, updated derivation for censoring subjects.	Section 2.6.2 and 2.11
			Corrected 'ratio of event rate' to 'hazard ratio'.	
			Added clarification all safety evaluations will be performed on the safety set (SAF), 'unless otherwise specified'.	Section 2.7
			Added exposure adjusted AE rates analysis.	Section 2.7.1
			Clarified ClinicalTrials.gov and EudraCT will be on treatment emergent adverse events.	
			List of AE of Special Interest updated.	Section 2.7.1.1
			Updated 'China specific safety analysis' to 'Other significant adverse events'.	Section 2.7.1.2
			For hepatic-related laboratory tests (newly occurring liver enzyme abnormalities), even if few subjects in a certain group of latent TB status, a summary table will display split by latent TB status.	Section 2.7.3
			Precision added in case there is some missing daily scores.	Section 2.7.4.4
			[redacted]	[redacted]
			[redacted]	
			Corrected first four secondary objectives to first three.	Section 3

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			Corrected SAF definition and added updated baseline definition, updated concomitant medications analysis [redacted]	Section 4
			Removed * and footnote, since study week is derived from study day, regardless of when study visits actually occur.	Section 5.1
			Update on DLQI in order to use Visit (not study week).	
			Removed date imputation rules, since these are specified in the Programming Dataset Specifications.	
			Precisions added for some analysis.	Section 5.6.1.1 Section 5.6.2.2 Section 5.6.2.3 Section 5.6.2.5 Section 5.6.2.7
			Added estimate and 95% CI to Multiple testing strategy.	Section 5.6.2.6
			G matrix updated and pre-processing of p-values added in the multiple testing strategy.	
			List of Protocol Deviations updated.	Section 5.7
			Removed sensitivity analysis of patients with incorrect TB status, since this is not required.	Section 5.9

Table of contents

List of abbreviations

AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Transaminase
ANCOVA	Analysis of covariance
AST	Aspartate Transaminase
ATC	Anatomical Therapeutic Classification
BMI	Body Mass Index
BOCF	Baseline Observation Carried Forward
CDFA	China Food and Drug Administration
CI	Confidence Interval
CSR	Clinical Study report
CSU	Chronic Spontaneous Urticaria
DLQI	Dermatology Life Quality Index
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
eCRS	Electronic Case Retrieval Strategy
eDiary	Electronic Diary
FAS	Full Analysis Set
GPS	Novartis Global Programming and Statistical environment
H1AH	H1 Antihistamines
HR	Hazard Ratio
IgE	Immunoglobulin E
IRT	Interactive Response Technology
ISS	Itch Severity Score
ISS7	Weekly Itch Severity Score
J2R	Jump to Reference
LHS7	Weekly size of largest hive score
LOCF	last observation carried forward
LSM	Least Squares Means
LTRA	Leukotrine Receptor Antagonist
MAR	Missing At Random
MedDRA	Medical Dictionary for Drug Regulatory Affairs
MID	Minimally Important Difference
MMRM	Mixed Effect Linear Model With Repeated Measures
MNAR	Missing Not At Random
NHS	Number of Hives Score
NHS7	Weekly Number of Hives Score
NovDTD	Novartis Drug and Therapy Dictionary
OR	Odds Ratio
[redacted]	[redacted]
[redacted]	[redacted]
PPS	Per-Protocol Set

PRO	Patient-reported Outcomes
RAN	Randomized Set
SAE	Serious Adverse Event
SAF	Safety Set
SAP	Statistical Analysis Plan
SD	Standard Deviation
SMQ	Standardized MedDRA query
SOC	System Organ Class
TB	Tuberculosis
TFL	Table, Figure, Listing
UAS	Urticaria Activity Score
UAS7	Weekly Urticaria Activity Score
ULN	Upper limit of normal range

1 Introduction

The purpose of Statistical Analysis Plan (SAP) is to describe the implementation of the statistical analysis planned in the protocol. The analysis planned in the SAP will be conducted on all subject data at the time the trial ends and the result will be described in the final Clinical Study Report (CSR). Protocol version 00 dated August, 10th 2016 has been referenced at the time of finalization of the SAP.

1.1 Study design

The study is a Phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study to evaluate the efficacy and safety of omalizumab administered subcutaneously as an add-on therapy for the treatment of patients aged 18-75 years with the diagnosis of refractory CSU and who remain symptomatic despite approved-dosed H1AH treatment. Patients will be randomized into three treatment arms (omalizumab 300 mg s.c., omalizumab 150 mg s.c., and placebo) in a 2:2:1 ratio, stratified by latent TB status at baseline. The study will consist of three epochs over 24 weeks:

- Screening epoch: Day -28 to Day -1
- Randomized-treatment epoch: Day 1 to Week 12
- Post-treatment follow-up epoch: Week 12 to Week 20

Approximately 420 patients will be enrolled at approximately 30 study sites.

The primary analysis time point is when the final analysis is conducted at the end of the study. There is no interim analysis planned in the study.

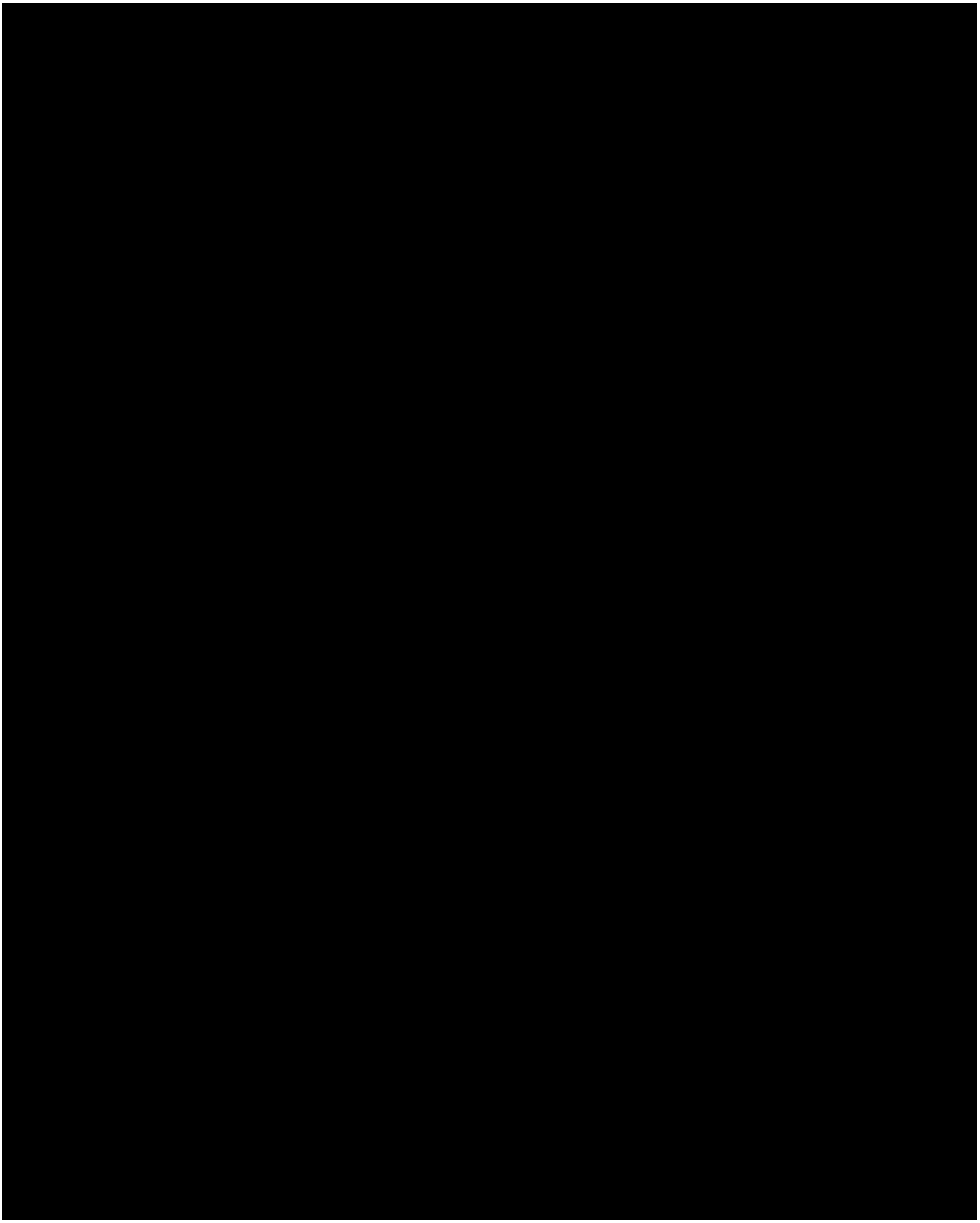
1.2 Study objectives and endpoints

Table 1-1 Objectives and related endpoints

Objective	Endpoint	Analysis
Primary		
To demonstrate the superiority of omalizumab 300 mg or 150 mg administered subcutaneously every 4 weeks in patients with refractory CSU receiving concomitant H1AH therapy with respect to change from baseline in weekly itch severity score (ISS7) at Week 12, compared to placebo	The primary efficacy variable is defined as the change from baseline of the ISS7 score after 12 weeks of treatment	Section 2.5
Secondary		
To demonstrate that patients with refractory CSU receiving concomitant H1AH who are treated with omalizumab 300 mg or 150 mg have a greater reduction from baseline in weekly urticaria activity score (UAS7) at Week 12, compared to placebo-treated patients To demonstrate that patients with refractory CSU receiving concomitant H1AH who are treated with omalizumab 300 mg or 150 mg have a greater	Variables and timepoint: UAS7 • Change from Baseline of UAS7 score after 12 weeks of treatment • Percentage of patients with	Section 2.6

Objective	Endpoint	Analysis
reduction from baseline in weekly number of hives score (NHS7) at Week 12 relative to placebo-treated patients To demonstrate that a greater percentage of patients with refractory CSU receiving concomitant H1AH who are treated with omalizumab 300 mg or 150 mg have UAS7 $\leq$ 6 at Week 12 relative to placebo-treated patients To demonstrate that a greater percentage of patients with refractory CSU receiving concomitant H1AH who are treated with omalizumab 300 mg or 150 mg achieve UAS7 = 0 at Week 12 relative to placebo-treated patients To demonstrate that a greater percentage of patients with refractory CSU receiving concomitant H1AH who are treated with omalizumab 300 mg or 150 mg achieve ISS7 Minimally Important Difference (MID) response at Week 12 relative to placebo-treated patients To demonstrate that patients with refractory CSU receiving concomitant H1AH who are treated with omalizumab 300 mg or 150 mg have a greater reduction from baseline in Dermatology Life Quality Index (DLQI) at Week 12 relative to placebo-treated patients To evaluate the efficacy of omalizumab compared with placebo in patients with refractory CSU receiving concomitant H1AH therapy with regards to time to ISS7 MID response by Week 12 To evaluate the safety of omalizumab compared with placebo in patients with refractory CSU receiving concomitant H1AH therapy with regards to the incidence and severity of adverse events and serious adverse events, vital signs and clinical laboratory evaluation at the end of the study	UAS7≤ 6 at Week 12 • Percentage of patients with UAS7=0 at Week 12 NHS7 • Change from Baseline of NHS7 score after 12 weeks of treatment ISS7 • Percentage of patients with ISS7 MID at week 12 • Time to ISS7 MID response by Week 12 DLQI • Change from Baseline of DLQI score after 12 weeks of treatment Safety • Percentage of patients with AE, with SAE, and who discontinue due to an AE • Exposure adjusted AE event rates • Percentage of patients with a clinically notable abnormality in Lab, ECG, and vital signs • Change from baseline in Lab, ECG, and vital signs	

[redacted]



2 Statistical methods

2.1 Data analysis general information

Data will be analyzed [redacted] (Contracted Clinical Research Organization) according to the data analysis Section 9 of the study protocol which is available in Appendix 16.1.1 of the CSR. Important information is given in the following sections and details are provided, as applicable, in Appendix 16.1.9 of the CSR.

SAS 9.4 will be used for generating study outputs used for clinical reports.

Unless otherwise stated, summary tables/figures/listings will be on all subjects included in the population under consideration. Data will be summarized with respect to demographic and baseline disease characteristics, efficacy, and safety assessments. The analysis will be conducted on all subject data at the time the trial ends.

The stratification factor latent TB status at baseline will be included in subgroup analysis where appropriate.

2.1.1 General definitions

Study drug/treatment refers to the investigational drug Omalizumab and placebo.

Date of first administration of study drug/treatment, or first date of study drug/treatment, refers to the date when the first dose of assigned treatment is administered. Date of last administration of study drug/treatment, or last date of study drug/treatment, refers to the date when the last dose of assigned treatment is administered.

Duration of exposure is defined as last date of study drug/treatment minus the first date of study drug/treatment plus 4 weeks (28 days), except if the subject died before the date of the last treatment + 27 days, then duration of exposure in days is defined as date of death minus the date of first study drug administration + 1 day.

Study day is defined as: if after first date of study drug, then Study Day = Date – first date of study drug + 1; if before first date of study drug, then Study Day = Date – first date of study drug; Study Day 1 is defined as the first date of study drug; there is no Study Day 0.

Baseline is defined as the last non-missing result before or on first date of study drug. If a subject did not take any study drug, then the baseline is defined as the last non-missing result before or on randomization date.

2.2 Analysis sets

Screened set: All subjects who signed the informed consent.

Randomized population (RAN): The RAN will include all randomized subjects, regardless of whether they took any study medication. Patients in RAN will be analyzed according to the treatment assigned at randomization.

Full analysis set (FAS): The FAS will include all randomized subjects who receive at least one dose of study drug with the exception of those who have inadvertently been randomized into the study. Subjects who are randomized due to erroneous use of the IRT system (identified as

being randomized, having no dose administration records and who discontinue the study immediately) will be excluded. Following the intent-to-treat principle, subjects will be analyzed according to the treatment they are assigned to at randomization. This analysis population will be used for all efficacy analyses unless otherwise specified.

Per-protocol set (PPS): The PPS will include all patients in the FAS who have received 3 doses of study drug during the treatment phase and without any major protocol deviations impacting the primary endpoint. Subjects will be analyzed according to the actual treatment received during the study as follows:

- Placebo: Patients who received only placebo injections (i.e., no active treatment) during the study
- 150 mg omalizumab: Patients who received at least one 150 mg omalizumab injection but no higher active dose level (i.e., 300 mg) during the study
- 300 mg omalizumab: Patients who received at least two 150 mg omalizumab injections at once during the study

This supplementary efficacy population will be used to assess the robustness of the primary analysis results. For the PPS, major protocol deviations impacting the primary endpoint will be considered for the exclusion of subjects and are reported in the section 5.7 Rule of exclusion criteria of analysis sets.

Safety set (SAF): The SAF consists of all subjects who take at least one dose of study medication. Subjects will be analyzed according to the actual treatment received during the randomized treatment epoch, as follows:

- Placebo: Patients who received only placebo injections (i.e., no active treatment) during the randomized treatment epoch
- 150 mg omalizumab: Patients who received at least one 150 mg omalizumab injection but no higher active dose level (i.e., 300 mg) during the randomized treatment epoch
- 300 mg omalizumab: Patients who received at least two 150 mg omalizumab injections at once during the randomized treatment epoch.

FAS, PPS, SAF: The protocol deviation codes leading to exclusion from the analysis sets defined above are presented in the section 5.7 Rule of exclusion criteria of analysis sets.

2.2.1 Subgroup of interest

Subgroup analyses will be performed by categories of the following demographic and baseline variable categories to evaluate the consistency of the primary efficacy results and key safety results. Summary analyses and the primary model will be applied repeatedly by the category of subgroup. Unless specified otherwise:

Subgroup analysis for efficacy will be performed by:

- Gender (male vs female)
- Age group (<65, >=65 years)
- Body weight group (<40, 40-<80, >=80 kg)
- Duration of CSU (<2, 2-10, >10 years at the time of Visit 1)

- Previous use of systemic treatment (including steroids and other systemic treatment including cyclosporin) for CSU (Yes vs No)
- Baseline ISS7 (<median, >=median)
- Baseline UAS7 (<median, >=median)
- Baseline total IgE (<median, >=median)
- Baseline presence of angioedema (Yes vs No)

These subgroup analysis will be used for the primary analysis (change from baseline in ISS7 at Week 12) and repeated also for UAS7 at week 12 and for NHS7 at week 12 which will use the same statistical method (MMRM model).

Subgroup analysis for safety will be performed by:

- Gender (male vs female)
- Age group (<65, >=65 years)
- Body weight group (<40, 40-<80, >=80 kg)
- Previous use of systemic treatment (including steroids and other systemic treatment including cyclosporin) for CSU (Yes vs No)
- Baseline total IgE (<median, >=median)

These subgroup analysis will be used for AEs summaries.

2.3 Patient disposition, demographics and other baseline characteristics

The analysis will be based on the RAN, unless otherwise specified.

2.3.1 Patient disposition

The number of screened subjects who completed the screening epoch will be given and the reasons for not entering the randomized treatment epoch will be summarized. The number and percentage of subjects in the RAN who completed/discontinued the treatment epoch, and the reason for discontinuation will be presented by treatment group. The number and percentage of subjects in the RAN who completed/discontinued the follow-up epoch, and the reason for discontinuation will also be presented by treatment group. Patient disposition during the treatment and follow-up epochs will be listed.

2.3.2 Demographic and baseline characteristics

Summary statistics will be presented for continuous demographic and baseline characteristic variables for each treatment group. The number and percentage of patients in each category will be presented for categorical variables for each treatment group and all patients.

The randomized treatment groups will be summarized by the following demographic variables:

- Gender (Male, Female)
- Age (years)
- Age group (< 65, >= 65),

- Ethnicity
- Race (Caucasian, Black, Asian, Native American, Pacific Islander, Unknown, Other)
- Weight (kg)
- Weight group (<40, 40 - < 80, >= 80)
- Height (cm)
- Body Mass Index (BMI): (kg/m²)– calculated as weight (kg) / (height (m))^2
- BMI group (<25, 25 - < 30, >=30)

Baseline disease characteristics will also be summarized for the following variables, including:

- Duration of CSU (years)
- Duration of CSU (<2, 2-10, > 10)
- Previous number of CSU medications in classes (<= 3, >3)
- Previous use of systemic treatment (including steroids and other systemic treatment including cyclosporin) for CSU (Yes, No). The list of systemic treatment is provided in the section 5.8.
- Total IgE level (ng/mL)
- Free IgE level (ng/mL)
- Latent TB status (Yes, No) from laboratory data. Further information are provided in the section 5.9.
- In-clinic UAS
- Weekly urticaria activity score (UAS7)
- Weekly itch severity score (ISS7)
- Weekly number of hives score (NHS7)
- Presence of angioedema (Yes, No)

2.3.3 Medical history

Any condition entered on the Medical history eCRF will be coded using the MedDRA dictionary. The number of patients with medical history will be summarized by primary system organ class, preferred term and treatment group for the RAN. Protocol solicited medical history will be also summarized.

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

The analysis of study treatment data will be based on the SAF, unless otherwise specified.

2.4.1 Study treatment / compliance

The duration of exposure to study treatment will be summarized by treatment group. In addition, the number of doses and total cumulative dosage will be presented.

Duration of exposure in days is defined as the date of the last treatment minus the date of first study drug administration plus 4 weeks (28 days), except if the subject died before the date of the last treatment + 27 days, then duration of exposure in days is defined as date of death minus the date of first study drug administration + 1 day.

2.4.2 Prior, concomitant and post therapies

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Any medication given at least once between the day of first dose of randomized study treatment and the date of the last study visit will be a concomitant medication, including those which were started pre-baseline and continued into the period where study treatment is administered.

Prior and concomitant medications will be identified using Novartis Drug and Therapy Dictionary (NovDTD) including Anatomical Therapeutic Chemical (ATC) code.

Concomitant medications will be summarized by treatment separately for CSU related and non-CSU related medications. CSU-related medications will be summarized by ATC class and preferred term. Non-CSU related concomitant medications will be summarized by the ATC class and preferred term. Concomitant medications will be summarized by epoch; the randomized treatment epoch and newly onset during post-treatment follow-up epoch.

Prior medications not for CSU will be summarized by ATC class, preferred term and treatment group. Urticaria therapy prior to screening (entered on the *Prior urticaria therapy* eCRF page) will be summarized by type of therapy, preferred term and treatment group.

Significant surgery and medical procedures will be summarized by primary system organ class and MedDRA preferred term.

Rescue medication will be listed by treatment group.

2.5 Analysis of the primary objective

Analyses will be based on the patients in FAS, unless otherwise specified.

2.5.1 Primary endpoint

The primary efficacy variable is change from baseline in weekly itch severity score (ISS7, a component of the UAS7, see protocol Section 6.4.1.1) at Week 12.

The daily itch score is the average of the morning and evening itch severity scores. The baseline ISS7 is the sum of the daily itch severity scores over the 7 days prior to the first study drug administration, and the ISS7 at Week 12 is the sum of daily itch scores over the 7 days during study days 78 – 84 (prior to the planned day of the Week 12 visit, study day 85). The same principles of calculating baseline and Week 12 weekly scores will be applied to each weekly outcome unless otherwise stated. See section 5.1 for details.

2.5.2 Statistical hypothesis, model, and method of analysis

A mixed-effect linear model with repeated measures (MMRM) will be used to obtain the least squares mean (LSM) estimate for each treatment group for change from baseline in ISS7 at Week 12 (the primary endpoint). The MMRM model will include terms of treatment group, week (1 to 12), baseline score, baseline score-by-week interaction, and treatment-by-week interaction as fixed effects. Treatment group and week will be fitted as categorical variables, and baseline score as a continuous covariate. The within-patient correlation will be modeled using the unstructured covariance matrix ([Mallinckrodt et al 2001]). If the model does not

converge, the compound symmetry covariance structure will be used. The occurrence of missing data will be assumed to be missing at random (MAR).

The difference in LSM estimates between treatment groups, together with a 95% CI, will be presented.

A multiplicity adjustment for two primary efficacy comparisons based on above model will be made according to the overall study testing strategy provided in Section 2.6.2. The statistical analyses will test the null hypothesis of no difference between the placebo and each omalizumab dose group.

Even if latent TB status at baseline was a stratification variable for the randomization, it will not be included in the efficacy model. Indeed, this variable will have no influence on the efficacy analysis.

2.5.3 Handling of missing values/censoring/discontinuations

The primary endpoint will be analyzed using an MMRM which is valid under the MAR (Missing-At-Random) assumption ([Rubin 1976]). Patients who have at least one post-baseline are included in the MMRM analysis. This includes not only the patients who have completed 12 weeks of treatment (with a complete 12-week longitudinal data), but also those who discontinue from study treatment early (although only able to contribute to a partial time profile). This is under the assumption that dropouts would follow the similar data pattern like other patients who complete the treatment period in the same treatment group, as if they had not discontinued from the study treatment.

Patients who discontinue from study treatment early (have less than three doses of study drug) will remain in the study and follow the procedures described in protocol Section 5.6.2. Nevertheless, the data on and post the date of the last treatment + 4 weeks (28 days) will be treated as missing in the primary analysis. This principle will be applied on all parameters from UAS questionnaire: ISS, UAS, NHS for the MMRM model.

The ISS7 is the sum of the average daily itch scores over 7 days each week. The daily itch scores are calculated based on daily eDiary entries for itch. The daily itch score is then calculated as the average of the morning and evening itch scores.

When either the morning or evening score is missing, the non-missing itch score for that day (morning or evening) will be used as the daily score. When one or more of the daily itch scores are missing, the following principles will be applied to handle the missing data:

- If a patient has at least 4 non-missing daily itch scores within the 7 days in a week, the ISS7 score is calculated as the sum of the available eDiary itch scores in that week, divided by the number of days that have a non-missing diary itch score, multiplied by 7.
- If there are less than 4 non-missing daily itch scores within the 7 days in a week, then the ISS7 score is missing for the week.

2.5.4 Supportive analyses

The primary efficacy analysis will be repeated on the PPS to assess the robustness of the primary result.

The impact of missing data on the primary analysis results will be assessed by repeating the analysis using different missing data assumptions to handle missing data. The supportive analyses may include:

- An analysis of covariance (ANCOVA) model with missing Week 12 itch scores imputed by carrying forward the patients' baseline scores (BOCF). The ANCOVA model would include treatment group as a factor, and baseline score as a covariate.
- An ANCOVA model with missing Week 12 itch scores imputed by carrying forward the patients' last non-missing weekly score (LOCF). The ANCOVA model would include treatment group as a factor, and baseline score as a covariate.
- Jump-to-Reference (J2R) multiple imputation (control-based pattern imputation) approach

2.6 Analysis of secondary efficacy objective(s)

All of the secondary efficacy variables will be analyzed using the FAS unless otherwise specified.

2.6.1 Secondary endpoints

The secondary endpoints are described below:

UAS7

- Change from Baseline in UAS7 score at Week 12
- Percentage of patients with UAS7≤ 6 at Week 12
- Percentage of patients with UAS7=0 at Week 12

NHS7

- Change from Baseline in NHS7 score at Week 12

ISS7

- Percentage of patients with ISS7 MID at week 12
- Time to ISS7 MID by Week 12

DLQI

- Change from Baseline in overall DLQI score at Week 12

Analysis of secondary endpoints is described below.

2.6.2 Statistical hypothesis, model, and method of analysis

Testing strategy

The following primary and secondary hypotheses will be included in the testing strategy. To ensure the family-wise type I error rate (α) is kept at an overall level of less than 5%, a flexible gate-keeping procedure ([Bretz et al 2009]) will be employed as described below. The procedure allows the type-one error rate associated with a rejected hypothesis to be reallocated among the remaining (un-rejected) hypothesis tests according to a set of pre-specified rules. The hypotheses are organized to describe the order in which different sets of hypotheses will be tested.

Primary objectives (as described in Section 2.5):

H_1: Omalizumab 300 mg is not different to placebo with respect to change from baseline in ISS7 (weekly itch severity score) at Week 12

H_2: Omalizumab 150 mg is not different to placebo with respect to change from baseline in ISS7 (weekly itch severity score) at Week 12

Secondary objectives:

H_3: Omalizumab 300 mg is not different to placebo with respect to change from baseline in UAS7 (weekly urticaria activity score) at Week 12

H_4: Omalizumab 150 mg is not different to placebo with respect to change from baseline in UAS7 (weekly urticaria activity score) at Week 12

H_5: Omalizumab 300 mg is not different to placebo with respect to change from baseline in NHS7 (weekly number of hives score) at Week 12

H_6: Omalizumab 150 mg is not different to placebo with respect to change from baseline in NHS7 (weekly number of hives score) at Week 12

H_7: Omalizumab 300 mg is not different to placebo with respect to the percentage of patients with UAS7 $\leq$ 6 response at Week 12

H_8: Omalizumab 150 mg is not different to placebo with respect to the percentage of patients with UAS7 $\leq$ 6 response at Week 12

H_9: Omalizumab 300 mg is not different to placebo with respect to the percentage of patients with UAS7 = 0 response at Week 12

H_{10}: Omalizumab 150 mg is not different to placebo with respect to the percentage of patients with UAS7 = 0 response at Week 12

H_{11}: Omalizumab 300 mg is not different to placebo with respect to the percentage of patients with ISS7 MID response at Week 12

H_{12}: Omalizumab 150 mg is not different to placebo with respect to the percentage of patients with ISS7 MID response at Week 12

H_{13}: Omalizumab 300 mg is not different to placebo with respect to change from baseline in overall DLQI score at Week 12

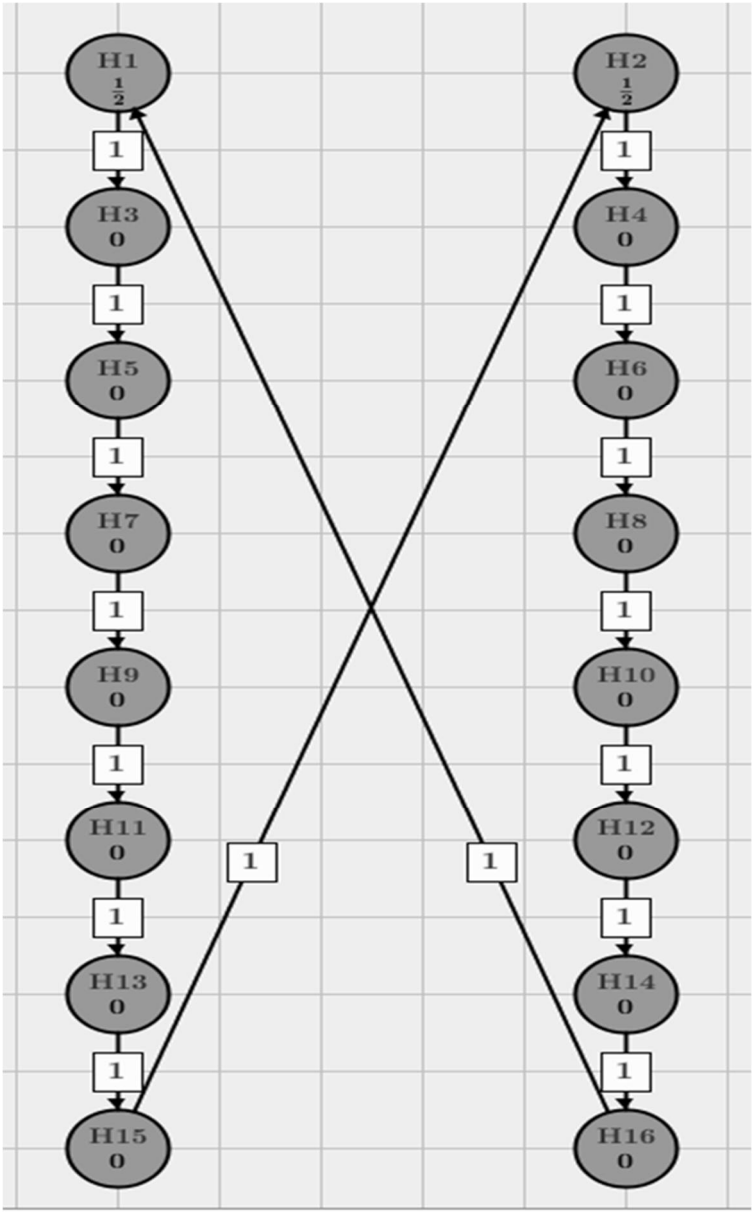
H_{14}: Omalizumab 150 mg is not different to placebo with respect to change from baseline in overall DLQI score at Week 12

H_{15}: Omalizumab 300 mg is not different to placebo with respect to the time to ISS7 MID by Week 12

H_{16}: Omalizumab 150 mg is not different to placebo with respect to the time to ISS7 MID by Week 12

The graphical approach of ([Bretz et al 2009]) for sequentially rejective testing procedures is used to illustrate the testing strategy in Figure 2-1.

Figure 2-1 Testing strategy for primary and secondary endpoints



Note: As shown in the above graph, initially each H_1 and H_2 will be assigned $\alpha/2$ to test the individual hypotheses simultaneously and move forward.

The family-wise error rate will be set to $\alpha = 5\%$ (2-sided) and will be controlled using the proposed hierarchical testing strategy as illustrated in Figure 2-1.

First, each of the hypotheses (H_1 and H_2) for the primary objective (based on change from baseline in ISS7 at Week 12) for omalizumab 300 mg and 150 mg versus placebo will be tested simultaneously at $\alpha/2$.

If at least one of H_1 and/or H_2 is rejected, then H_3 and/or H_4, respectively, will be tested at $\alpha/2$. If at least one of H_3 and/or H_4 is rejected, then H_5 and/or H_6, will be tested, respectively. A similar process applies until H_{15} and H_{16}. Once all hypotheses for an omalizumab dose are rejected, then the respective $\alpha/2$ can be passed on to the other dose's hypotheses, if they are not already rejected at $\alpha/2$. In the description above, rejection of a hypothesis refers to rejection of the two-sided hypothesis; however the significance level of a rejected hypothesis is only passed on according to the graphical procedure for the test of another hypothesis if the treatment effect is in favor of omalizumab.

Change from baseline in weekly urticarial activity score (UAS7) at Week 12

The urticaria activity score (UAS) is a composite score (itch severity score and number of hives score) described in protocol Section 6.4.1.1. For each of the morning and evening UAS score, it is calculated as the sum of the itch severity score and number of hives score according to eDiary entries. The daily UAS is the average of the morning and evening UAS scores, and UAS7 is the sum of daily UAS scores over 7 days.

The missing data will be handled in the same way as described in Section 2.5.3.

Treatment comparisons of 300 mg vs placebo (H_3) and 150 mg vs placebo (H_4) in change from baseline to Week 12 in the UAS7 will be made using an MMRM model with similar terms as the primary analysis but baseline UAS7 as a covariate (Refer to Section 2.5.2).

Change from baseline in weekly number of hives score (NHS7) at Week 12

The weekly number of hives score (NHS7) will be handled using the same principles as described for the primary endpoint in Section 2.5.1 and Section 2.5.3.

Treatment comparisons of 300 mg vs placebo (H_5) and 150 mg vs placebo (H_6) in change from baseline to Week 12 in the NHS7 will be made using an MMRM model with similar terms as the primary analysis but baseline NHS7 as a covariate (Refer to Section 2.5.2).

Percentage of patients with UAS7 ≤ 6 at Week 12

Treatment comparisons of 300 mg vs placebo (H_7) and 150 mg vs placebo (H_8) in the percentage of patients with UAS7 $\leq$ 6 at Week 12 will be made using a logistic regression model with treatment group as a factor and baseline UAS7 as a covariate.

A patient with missing data at Week 12 will be imputed as a responder if the patient was a responder at Week 10 and Week 11, otherwise as a non-responder.

Percentage of patients with UAS7 = 0 at Week 12

Treatment comparisons of 300 mg vs placebo (H_9) and 150 mg vs placebo (H_{10}) in the percentage of patients with UAS7 = 0 at Week 12 will be made using a logistic regression model with treatment group as a factor and baseline UAS7 as a covariate.

A patient with missing data at Week 12 will be imputed as a responder if the patient was a responder at Week 10 and Week 11, otherwise as a non-responder.

Percentage of patients with ISS7 MID response at Week 12

The ISS7 MID response is defined as a reduction from baseline in ISS7 of ≥ 5 points.

Treatment comparisons of 300 mg vs placebo (H_{11}) and 150 mg vs placebo (H_{12}) in the percentage of patients with ISS7 MID response at Week 12 will be made using a logistic regression model with treatment group as a factor and baseline ISS7 as a covariate.

A patient with missing data at Week 12 will be imputed as a responder if the patient was a responder at Week 10 and Week 11, otherwise as a non-responder.

Change from baseline in overall DLQI at Week 12

DLQI is a PRO instrument, described in protocol Section 6.4.2.1. An overall score will be calculated according to the scoring manual given in protocol Appendix 3. The baseline and up to Week 12 overall DLQI scores will be derived from the questionnaires assessed at the Day 1 and up to Week 12 visits.

Treatment comparisons of 300 mg vs placebo (H_{13}) and 150 mg vs placebo (H_{14}) in change from baseline to Week 12 in overall DLQI score will be made using an MMRM model with similar terms as primary analysis but baseline DLQI as a covariate (Refer to Section 2.5.2).

Time to ISS7 MID response by Week 12

The ISS7 MID response is defined as a reduction from baseline in ISS7 of ≥ 5 points. Time to ISS7 MID response is the time (in weeks) from the date of the first dose to the date where ISS7 MID response is first achieved during Week 1 to 12. If no ISS7 MID response is achieved by Week 12, then if the patient was a treatment completer (took 3 doses of study drug) they will be censored at the week of the last non-missing weekly score up to Week 12, and if the patient was not a treatment completer (took < 3 doses of study drug) they will be treated as censored at the maximum of (the last dose date + 28 days – 1)/7 and the last non-missing weekly score, to a maximum of 12 weeks.

Treatment comparisons of 300 mg vs placebo (H_{15}) and 150 mg vs placebo (H_{16}) will be performed using a Cox proportional hazard (PH) model with treatment group as a factor and baseline ISS7 as a covariate. The hazard ratio and 95% CI will be reported.

Kaplan-Meier analysis stratified by treatment group will be also presented with log-rank test and displayed graphically.

Subgroup analysis

All the sub-group analyses conducted for the primary efficacy variable (described in Section 2.2.1) will be repeated for UAS7 at week 12 and NHS7 (weekly number of hives score) at week 12. The primary model will be applied repeatedly by the category of subgroup.

2.7 Safety analyses

All safety evaluations will be performed on the safety set (SAF), unless otherwise specified.

2.7.1 Adverse events (AEs)

All the AEs occurring after providing written informed consent will be recorded on the Adverse Event eCRF page. AEs starting on or after the first dose of study treatment or events present prior to the first dose of study treatment but increased in severity based on preferred term will be classified as treatment emergent AEs. Overall AEs, SAEs, AEs by severity will be also summarized for each epoch (randomized-treatment and post-treatment follow-up) when necessary. Non-treatment emergent AEs (occurring after providing written informed consent but before first dose of study treatment) will not be summarized but listed only.

Treatment emergent AEs will be summarized by presenting, for each treatment group, the number and percentage of patients having any AE, having an AE in each primary system organ class and having each individual AE (preferred term). Summaries will be also presented for treatment emergent AEs by severity and for study treatment related AEs. If a patient reported more than one AE with the same preferred term, the AE with the greatest severity will be presented. If a patient reported more than one AE within the same primary system organ class, the patient will be counted only once with the greatest severity at the system organ class level, where applicable. Separate summaries will be provided for death, serious adverse event, and adverse events leading to discontinuation.

In addition, exposure adjusted AE rates (incidence rates) will be provided for each treatment group, by primary system organ class and preferred term.

Subgroup analysis will be performed for AE summaries, unless otherwise specified (see Section 2.2.1 for subgroup definition).

Clinicaltrials.gov and EudraCT

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on treatment emergent adverse events which are not serious adverse events with an incidence greater than 2% and on treatment emergent serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term 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 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
- 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 Adverse events of special interest / grouping of AEs

Treatment emergent AEs of special interest for omalizumab treatment will be also summarized. AEs of special interest for omalizumab treatment include the following, specified as compound-level risk factors defined in the electronic Case Retrieval Strategy (eCRS):

- Anaphylaxis - anaphylactoid reactions (narrow)
- Churg Strauss Syndrome - Hypereosinophilic syndrome
- Arterial Thromboembolic Events
- Malignant neoplasms

Summary tables that present numbers and percentages of patients with the AEs of special interest will be presented by standardized MedDRA Query (SMQ) (if applicable), preferred term and treatment. In addition, the hepatic disorders events will be tabulated by latent TB at baseline (Yes/No).

A sensitivity analysis will be done in order to summarize the hepatic disorders events among patients whose TB status may not be interpretable due to the incubation temperature violation (as described in Section 5.9). The current version of eCRS is stored in GPS folder.

2.7.1.2 Other significant adverse events

Treatment emergent AEs leading to treatment discontinuation will be summarized and listed. Treatment emergent AEs of special interest will be considered as treatment emergent other significant AEs as well.

2.7.2 Deaths

Overall death will be summarized and listed.

2.7.3 Laboratory data

Laboratory values that the laboratory reports to be below or above the limit of quantification will be imputed by the respective limit of quantification.

The summary of laboratory evaluations will be presented for three groups of laboratory tests (hematology, serum chemistry and urinalysis). Descriptive summary statistics for the change from baseline to each study visit 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 subjects with both baseline and post baseline values.

The notable criteria for platelet count ($\leq$ 100 x10E9/L) will be used to calculate number and percentage of patients with newly occurring or worsening notable abnormalities occurring while on study. A case will be considered as newly occurring if the value for a laboratory evaluation is not notable or missing at baseline but is notable thereafter; a case will be considered as worsening if the value for a laboratory evaluation is notable at baseline and at least one post-baseline value is worse than baseline.

The hepatic-related laboratory tests (newly occurring liver enzyme abnormalities) will also be tabulated by latent TB at baseline (Yes/No). A sensitivity analysis will be done to summarize the hepatic-related laboratory tests among patients whose TB status may not be interpretable due to the incubation temperature violation (as described in Section 5.9).

To evaluate potential drug-induced liver injury, newly occurring liver enzyme abnormalities at any time post-baseline will also be summarized while on study based on the event criteria given below:

- 3 x-, 5 x-, 8 x-, 10 x-, and 20 x upper limit of normal range (ULN) elevations of AST, 3 x-, 5 x-, 8 x-, 10 x-, and 20 x upper limit of normal range (ULN) elevations of ALT, and either AST or ALT
- elevation of TBL to > 1 x ULN, > 1.5 x ULN, and > 2 x ULN
- elevation of ALP to > 1.5 x ULN, > 2 x ULN, > 3 x ULN, and > 5 x ULN
- elevation of AST or ALT to >3 x ULN accompanied by elevated TBL (>1.5 x ULN, > 2 x ULN)
- elevation of ALP to (>3 x ULN, > 5 x ULN) accompanied by elevated TBL > 2 x ULN
- Potential Hy's Law: (AST or ALT > 3 x ULN) and TBL > 2 x ULN and ALP =< 2 x ULN

For a criterion with combined components except for Hy's Law, the elevations do not have to occur at the same post-baseline time point, which implies that the cases can be identified only by the highest post-baseline value. While for potential Hy's Law case, all the elevations must occur at the same post-baseline time point. A case will be considered as newly occurring if a criterion is not met or missing at baseline but is met thereafter.

2.7.4 Other safety data

2.7.4.1 Anti-omalizumab antibody

A summary of anti-omalizumab antibodies will be provided by treatment group.

2.7.4.2 ECG and cardiac imaging data

Not applicable.

2.7.4.3 Vital signs

Analysis of the vital sign measurements using summary statistics for the change from baseline for each post-baseline visit will be performed. These descriptive summaries will be presented by vital sign and treatment group. Change from baseline will only be summarized for subjects with both baseline and post-baseline values. Subjects with newly occurred notable vital signs as defined below will be listed.

- 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)

2.7.4.4 Other

Healthcare utilization (calling a doctor, nurse, or nurse practitioner) based on patient daily diary data will be summarized for each treatment group. In case of missing daily scores, observed data will be used.

For patients reporting angioedema, the action(s) taken in response to their angioedema based on patient daily diary data will be summarized for each treatment group.

[redacted]

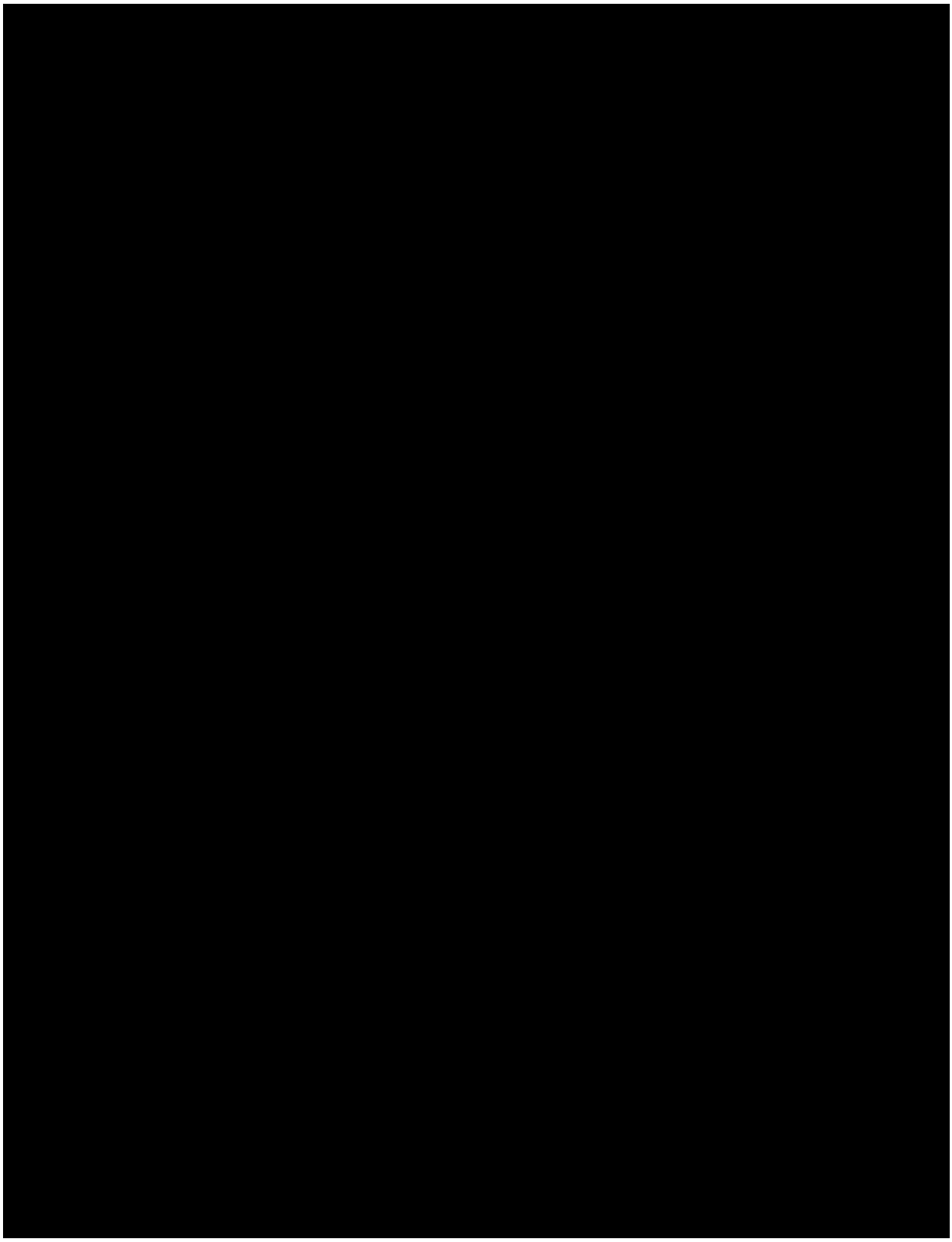
2.9 Patient-reported outcomes

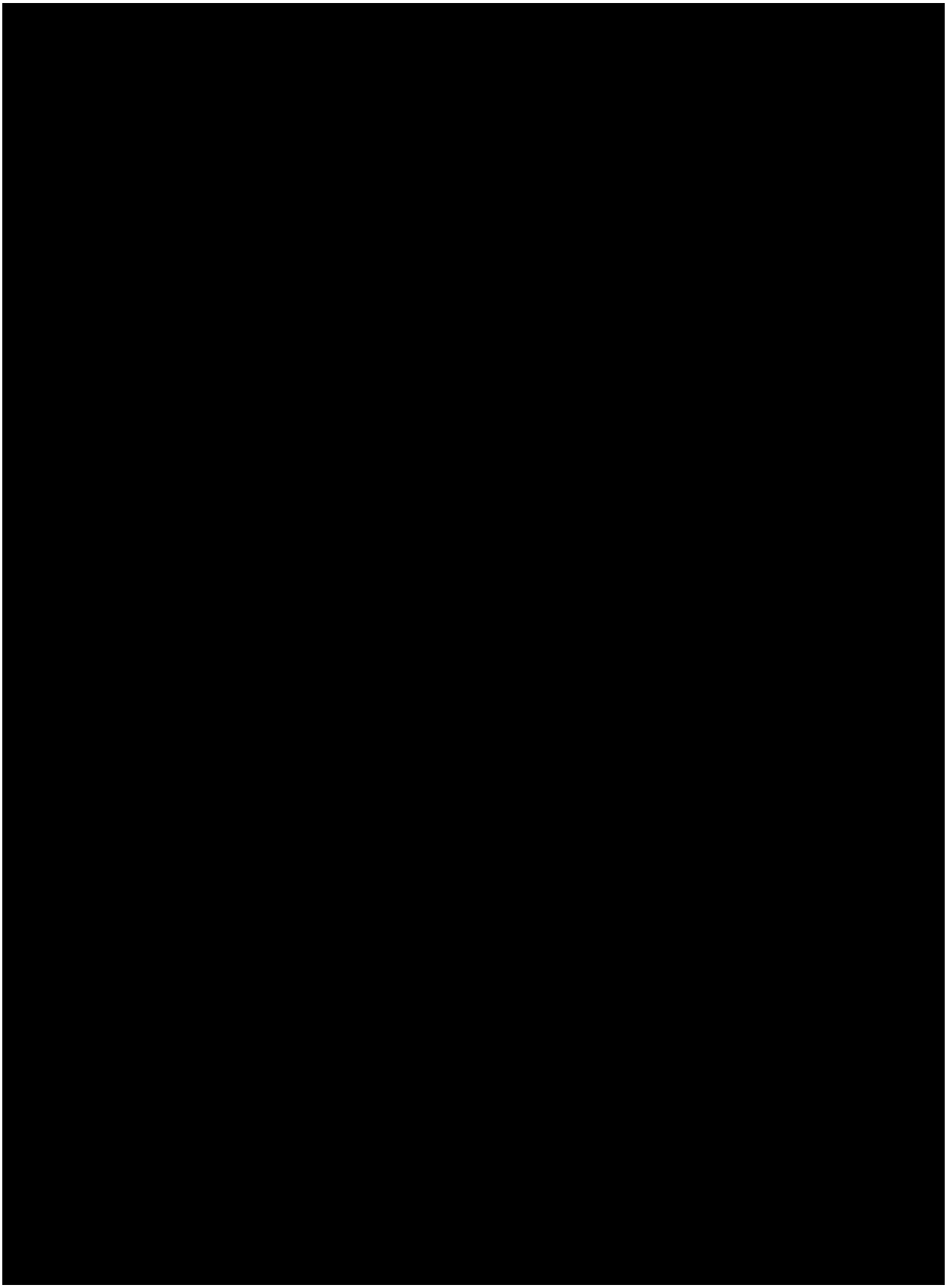
Not applicable.

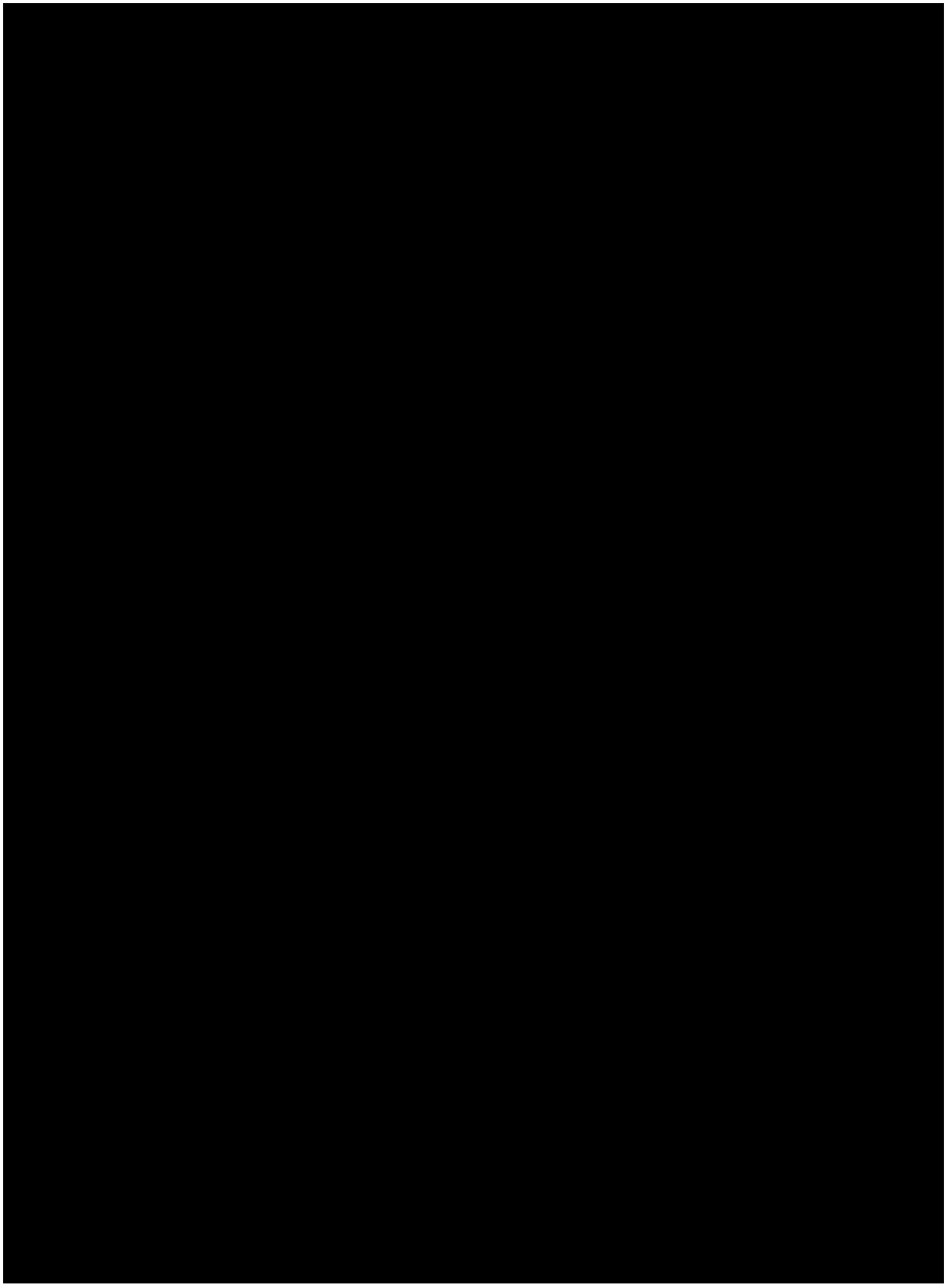
2.10 Biomarkers

Not applicable.

[redacted]







[redacted]

2.12 eDiary compliance analysis

To evaluate the compliance of eDiary completion, number of missing eDiary weekly scores of each patient will be summarized by treatment group and by study week. FAS will be used for this analysis.

2.13 Interim analysis

No interim analysis is planned in this study.

3 Sample size calculation

The study is sized to ensure sufficient power to demonstrate meaningful efficacy based on the data from the full analysis set. In addition, for registration of a biologic compound the China health authority (China Food and Drug Administration, CFDA) requires at least 300 patients to be treated in the test drug as the minimum for safety evaluation. Assuming 10% dropout rate by week 12, the study will randomize 420 patients in a ratio of 2:2:1 to omalizumab 300 mg, omalizumab 150 mg, and placebo to ensure 375 patients completing Week 12 with 300 patients in the two omalizumab dose groups.

The below power evaluation for the planned hypothesis tests is based on a sample size of 375 (2:2:1) patients who complete 12 weeks of treatment duration. The powers are examined according to the hierarchy order of the multiplicity type I error control plan with the overall alpha level controlled at 0.05 (2-sided) as outlined in Figure 2-1. It is assumed Study E2305 will repeat the performance of the previous pivotal studies (Q4881g, Q4882g, and E2306). Except for ethnicity, these 3 studies have a similar patient population with consistent results for the primary and secondary endpoints. There is no statistical evidence suggesting heterogeneity among the studies. A meta-analysis with fixed-effects model was performed based on summary data from these studies to combine the estimates of the treatment effect and standard deviation (SD). The obtained pooled estimates on the treatment effect size and SD are then used as the assumed values for the alternative hypotheses to evaluate the power for each endpoint. To understand the impact of the strength of dependency between endpoints on the power, the power was evaluated under 3 scenarios of correlation (0, 0.5, 0.9), i.e. completely independency, moderate dependency, and strong dependency. The results are summarized in Table 3-1.

For the primary endpoint (change from baseline weekly itch score at Week 12), the study will offer

- a power of > 99.9% to detect a difference of 4.73 between 300 mg and placebo with the assumed SD of 5.28
- a power of > 93% to detect a difference of 2.73 between 150 mg and placebo with the assumed SD of 5.55

The power was evaluated for the primary and secondary endpoints as a whole according to the testing strategy defined in Section 2.6.2 (Figure 2-1) using the gMCP package in the R software.

According to Table 3-1, the power evaluation is summarized as follows.

- The power for 300 mg is maintained at nearly 95% and above for all testing hypotheses regardless of the strength of dependency between endpoints.
- The power for 150 mg is at least 80% for the testing hypotheses up to the endpoint of percentage of patients with UAS7 ≤ 6 at Week 12. The power for the rest of endpoints is decreased to 65% and below, and can be as low as 18% if the dependency between endpoints is none.

By reviewing and using clinical data from previous studies in the power assessment for the targeted hypothesis tests for the current study, the total sample size of 420 (including 10% dropout) with 2:2:1 assignment ratio is considered appropriate to allow sufficient power to achieve the primary objective and all the first three secondary objectives for both 300 mg and 150 mg. The study also has high power to achieve the last four secondary objectives for 300 mg, although the power is < 65% and below for 150mg.

Table 3-1 Power analysis of primary and key secondary endpoints

Endpoint	Hypo-thesis	Comparison	Parameter assumptions[1]	Power[2]		
				Correlation between endpoints		
				None (0)	Moderate (0.5)	Strong (0.9)
Change from baseline in ISS7 at W12	H_1	300 mg vs Placebo	Δ=4.73 SD=5.28	>0.999	>0.999	>0.999
	H_2	150 mg vs Placebo	Δ=2.73 SD=5.55	0.933	0.931	0.932
Change from baseline in UAS7 at W12	H_3	300 mg vs Placebo	Δ=11.16 SD=11.54	>0.999	>0.999	>0.999
	H_4	150 mg vs Placebo	Δ=6.31 SD=11.84	0.899	0.906	0.925
Change from baseline in NHS7 at W12	H_5	300 mg vs Placebo	Δ=6.22 SD=6.82	>0.999	>0.999	>0.999
	H_6	150 mg vs Placebo	Δ=3.52 SD=6.77	0.859	0.880	0.917
% of patients with UAS7 ≤ 6 at W12	H_7	300 mg vs Placebo	Log OR=2.09 SD=2.48 (~ p_{300}=61%, p_0=16%)	>0.999	>0.999	>0.999
	H_8	150 mg vs Placebo	Log OR=1.32 SD=2.46 (~ p_{150}=41%, p_0=16%)	0.830	0.864	0.914
% of patients with UAS7 = 0 at W12	H_9	300 mg vs Placebo	Log OR=2.27 SD=3.46 (i.e. p_{300}=41%, p_0=7%)	0.992	0.992	0.991

Endpoint	Hypo-thesis	Comparison	Parameter assumptions[1]	Power[2]		
				Correlation between endpoints		
				None (0)	Moderate (0.5)	Strong (0.9)
	H_{10}	150 mg vs Placebo	Log OR=1.21 SD=3.63 (~ p_{150}=19%, p_0=7%)	0.540	0.616	0.654
% of patients with ISS7 MID at W12	H_{11}	300 mg vs Placebo	Log OR=1.57 SD=2.34 (~ p_{300}=81%, p_0=47%)	0.986	0.986	0.989
	H_{12}	150 mg vs Placebo	Log OR=0.78 SD=2.09 (~ p_{150}=65%, p_0=47%)	0.408	0.537	0.631
Change from baseline in Overall DLQI at W12	H_{13}	300 mg vs Placebo	Δ=3.55 SD=5.86	0.965	0.969	0.977
	H_{14}	150 mg vs Placebo	Δ=1.91 SD=6.25	0.237	0.404	0.536
Time to ISS7 MID	H_{15}	300 mg vs Placebo	Log HR=0.7 SD=1.14 (~ S_{300}=6%, S_0=25%)	0.947	0.955	0.970
	H_{16}	150 mg vs Placebo	Log HR=0.43 SD=1.15 (~ S_{150}=12%, S_0=25%)	0.179	0.372	0.530

[1] Parameter assumptions are based on a meta-analysis of Studies Q4881g, Q4882g, and E2306

Δ= Difference; OR=Odds Ratio; HR: Hazard Ratio

p_{300}, p_{150}, p_0: the probability of an event occurrence at Week 12 for Omalizumab 300mg, 150mg, and placebo

S_{300}, S_{150}, S_0: the probability of not having an event at and prior to Week 12 for Omalizumab 300mg, 150mg, and placebo

SD= Standard deviation for the data either normally distributed or on the scale with a normal approximation, e.g. logit(p) for binary data and log-log(S) for time-to-event data

[2] Power is calculated based on the total sample size of 420 assigned to treatment group 300 mg, 150 mg and placebo in ratio of 2:2:1 with 10% dropout rate during 12 weeks of treatment period, according to the hierarchy order of the multiplicity Type I error control scheme with overall alpha level controlled at 0.05 (2-sided).

4 Change to protocol specified analyses

Below changes from protocol specified analysis are made:

- Previous use of systemic treatments including steroids and other treatment (including cyclosporin) for CSU are summarized and used as subgroup analysis categories instead of previous use of systemic steroids for CSU.
- eDiary compliance analysis added to summarize the missing weekly eDiary scores.
- Weekly size of largest hive score (LHS7) is not summarize at baseline as not collected in this study.
- SAF definition updated in order to include all patients with at least one study medication intake.
- ECG will be not listed as these data are collected only in source document.
- [redacted]
- Baseline ISS7 definition updated in order to consider the first study drug administration (rather than the randomization day) since a subject can have the first study medication intake after the randomization date.
- CSU-related medications will be summarized by ATC class and preferred term. Also, non-CSU related concomitant medications will be summarized by ATC class and preferred term.
- [redacted]

5 Appendix

5.1 Study week definition

Definition of study week based on eDiary

A number of efficacy outcome measures (ISS7, NHS7, UAS7, [redacted]) are in the form of weekly scores derived from patient daily diary data. The study weeks are defined based on the study days starting with Day 1 (see Table 5-1), which is the day the patient receives the first study treatment. The study day for a particular diary date is calculated as [Date of diary] – [Date of first dose] + 1.

Table 5-1 Study week definitions based on study days

Study Week	Study Day	Study Week	Study Day
1	1-7	13	85-91
2	8-14	14	92-98
3	15-21	15	99-105
4	22-28	16	106-112

Study Week	Study Day	Study Week	Study Day
5	29-35	17	113-119
6	36-42	18	120-126
7	43-49	19	127-133
8	50-56	20	134-140
9	57-63		
10	64-70		
11	71-77		
12	78-84		

Furthermore, the baseline week is comprised of the 7 days prior to Day 1 (Day -7 to Day -1). Table 5-1 summarizes the time period (in study days) over which the data are used in the calculation of each weekly score. For most study weeks, weekly scores will be based on the 7 study days given in Table 5-1.

Definition of study week based on DLQI on visit date

For efficacy variables collected by visit (DLQI), the reported visit will be used in the analysis. No visit remapping will be applied.

5.2 Dermatology Life Quality Index (DLQI)

The DLQI measures functional disability of subjects with dermatological disorders that are 17 years or older of age and had been utilized as a relevant clinical measure in atopic dermatitis, as well as other dermatitis clinical trials. The DLQI is a simple, validated, self-administered 10-item questionnaire. The instrument contains six functional scales (i.e., symptoms and feeling, daily activities, leisure, work and school, personal relationships, treatment). For the DLQI, each question will be answered with the following response: “not at all”, “a little”, “a lot”, or “very much”. “Not relevant” is also a valid response. Seven scores will be derived from DLQI: the total score of each of the six dimensions as well as the total score over all items. The higher the score, the more quality of life is impaired.

The scoring of each question is as follows:

- Very much: Scored 3
- A lot: Scored 2
- A little: Scored 1
- Not at all: Scored 0
- Not relevant: Scored 0
- Question unanswered: Scored 0
- Question 7: “prevented work or studying”: Scored 3

The DLQI total score will be calculated by summing the score of each question resulting in a maximum of 30 and a minimum of 0. The higher the score, the more Quality of Life is impaired.

Meaning of DLQI Scores

- 0-1= no effect at all on subject's life
- 2-5= small effect on subject's life
- 6-10= moderate effect on subject's life
- 11-20= very large effect on subject's life
- 21-30= extremely large effect on subject's life

The DLQI will be analyzed under six headings as follows:

- Symptoms and feelings: question 1 and 2, score maximum 6
- Daily activities: question 3 and 4, score maximum 6
- Leisure: question 5 and 6, score maximum 6
- Work and school: question 7, score maximum 3
- Personal relationships: question 8 and 9: score maximum 6
- Treatment: question 10, score maximum 3

Interpretation of incorrectly completed questionnaires:

There is a very high success rate of accurate completion of the DLQI. However, sometimes subjects do make mistakes.

1. If one question is left unanswered this is scored 0.
2. If two or more questions are left unanswered the questionnaire will not be scored.
3. If question 7 is answered 'yes' this will be scored 3. If question 7 will be answered no or 'relevant' but then either 'a lot' or 'a little' is ticked this will then be scored 2 or 1.
4. If two or more response options are ticked, the response option with the highest score will be recorded.
5. If there is a response between two tick boxes, the lower of the two score options will be recorded.
6. If one item is missing from a two-item subscale, that subscale will not be scored.

Handling of missing values:

- If there is only one missing score per visit, it will be imputed with 0, and then the subscale including this item and the total score are derived accordingly.
- If there are two or more missing scores per visit, LOCF will be applied to the total score (i.e. LOCF is NOT applied to the 10 individual question scores for further derivation of the 6 subscale scores and 1 total score).

5.3 Handling of duplicate/mistaken data in eDiary

eDiary records are entered by subjects and if there are cases when more than one values are entered by subject for one entry due to technical reasons or by mistake, then the worst value will always be used in calculating weekly scores.

Systemic erroneous data due to device malfunction are always excluded from the analysis. The impact of such issue if occurs will be evaluated by Novartis team and additional analysis may be considered if necessary.

5.4 AEs coding/grading

Coding of AE will be done per MedDRA dictionary.

5.5 Laboratory parameters derivations

Not applicable.

5.6 Statistical models

5.6.1 Primary analysis

5.6.1.1 MMRM analysis

Change from baseline at Week 12 of ISS7 by treatment group is estimated using SAS procedure MIXED, and assuming unstructured covariance matrix. Compute least square means for change from baseline at Week 12 for each treatment group, and p-values are given with null hypothesis that there is no difference between active treatment group and placebo group.

The following code will be used for this analysis:

```
ods output Lsmeans=lsmeans_est
           Diffs=diff
           ConvergenceStatus=conv;
proc mixed data=dataset order=internal;
  class subject treatment week;
  model <change from baseline> = treatment week baseline treatment*week
baseline*week / s ddfm=kr;
  repeated week / subject = subject type = un;
  lsmeans treatment*week / cl pdiff;
run;

where change from baseline = change from baseline for the analysis parameter
       baseline = baseline score of the analysis parameter
       week = avisitn (from week 1 up to week 12 for Week 12 analysis / from week 1
up to week 20 for Week 20 analysis)
       treatment = treatment group assigned in numeric (1 = 300 mg, 2 = 150 mg, 3 =
Placebo)
       subject = usubjid
```

Note that the diff dataset created will include all the treatment interactions for each week. The estimates for the right weeks, e.g. Week 12, and right treatment comparisons, e.g. 300 mg vs Placebo, should be subsetted by taking “week = 12 and _week=12” and “treatment=1 and _treatment=3”.

If the model with an unstructured covariance matrix does not converge, SAS will give a warning as “Unable to make hessian positive definite” or “Unable to Converge”. In this case, the compound-symmetry structure should be used instead: `type=cs` will be used in the above code.

5.6.2 Other analysis

MMRM analysis are similar as defined for the primary analysis.

5.6.2.1 Analysis of covariance (ANCOVA) model

As a supportive analysis, change from baseline at Week 12 of ISS7 will be also estimated using analysis of covariance (ANCOVA) model using SAS procedure MIXED. Least square means for each treatment group are estimated, with p-values and confidence intervals generated under null hypothesis that no difference between active treatment group and placebo group.

Baseline observation carried forward (BOCF) and last observation carried forward (LOCF) will be used separately to impute missing data at Week 12.

The following code will be used for this analysis:

```
ods output LSMeans=lsmeans_est
           Estimates=diff_est;
proc mixed data=dataset order=internal;
  class treatment;
  model <change from baseline> = treatment baseline / s ddfm=kr;
  lsmeans treatment / cl pdiff;
  estimate "150 mg - Placebo" treatment 0 1 -1 / cl;
  estimate "300 mg - Placebo" treatment 1 0 -1 / cl;
run;

where change from baseline = change from baseline for the analysis parameter
       baseline = baseline score of the analysis parameter
       treatment = treatment group assigned in numeric (1 = 300 mg, 2 = 150 mg, 3 =
Placebo)
```

5.6.2.2 Jump-to-reference (J2R) multiple imputation (control-based pattern imputation) for missing data

A multiple imputation for missing data approach will be performed to check the sensitivity of the primary analysis. As MMRM performed in the primary analysis is under MAR assumption, Jump-to-reference (J2R) multiple imputation also known as control-based pattern imputation under MNAR assumption for missing data will be performed.

As the first step, SAS procedure MI is used to create a number of complete datasets using control-based pattern imputation. In step 2, each of these datasets is analyzed by the standard procedure (here it is PROC MIXED for MMRM model) and in step 3, the results of the multiple analyses are combined into a single analysis by using the SAS procedure MIANALYZE.

The input dataset should have one record per subject with baseline score as well as all available ISS7 scores from Week 1 up to Week 12. The output dataset will contain a pre-specified number of concatenated imputed datasets and a new variable that identifies the imputation index.

Once the imputation step is complete, the concatenated dataset can in turn be transposed back to the format with multiple records (weeks) per subject and then analyzed using PROC MIXED originally described in Section 5.6.1.1 with the additional by-variable of the index variable. The treatment differences (estimates and standard errors for 150 mg vs Placebo, 300 mg vs Placebo and 150 mg vs 300 mg) at week 12 for each imputed dataset should be saved in a dataset for further analysis using PROC MIANALYZE.

The imputation will be done in 2 steps:

1. To impute non-monotone missing data patterns – MAR assumption: in order to impute missing intermittent values.
2. To perform the control-based imputation – MNAR assumption: in order to impute the missing values after subject's discontinuation.

This method is based on the same methodology that was described by Ratitch, B. and O'Kelly, M. (2011).

The 1st step will be to impute all intermittent missing observations:
The following code will be used:

```
proc mi data=t_dat1b seed=1231 out=impdata_mono nimpute=1000;
   mcmc chain=multiple impute=monotone;
   var BASE _W1--_W12;
   by TRT01PN;
run;
where TRT01PN represents the planned treatment code (1: IGE 300 mg, 2: IGE 150 mg,
3: Placebo)
      _W1 to _W12 represents all post-baseline scores (from Week 1 up to week 12)
       BASE is the ISS7 score at baseline.
```

In this step, only the non-monotone missing data will be imputed by using Markov Chain Monte Carlo methodology. So, our dataset will be partially imputed.
Note that 1000 imputations will be used. This number is much larger than the default in most software packages. This high number of imputations decreases at the best the uncertainty of the imputation.

And then, 2nd step will be to impute the monotone missing observations by using the control-based imputation. This 2nd step will be done sequentially: one specific timepoint.

a) First, we will separate "impdata_mono" dataset in 2 datasets:
   - One dataset containing all subjects with active treatments and with no missing data at Week 1: impdata_mono_rest1
   - One dataset containing all subjects with active treatments with missing data at Week 1 and all subjects in the control group: impdata_mono_imp1

```
data impdata_mono_imp1 impdata_mono_rest1;
     set impdata_mono;
     if trt01pn in(1,2) and lastweek>=1 then output impdata_mono_rest1; /*subjects
taken active treatment and with non-missing values at Week 1*/
     else output impdata_mono_imp1; /*subjects taken active treatment and with
missing values at Week 1 or subjects in control group*/
run;
where TRT01PN represents the planned treatment code (1: IGE 300 mg, 2: IGE 150 mg,
3: Placebo)
      LASTWEEK is the last week with a non-missing score (defined prior to any
imputations).
```

b) Then, we will perform the imputation of missing data at Week 1 for the active treatments by using the control group:

```
proc sort data=impdata_mono_imp1; by _Imputation_ usubjid; run;
```

```
proc mi data=impdata_mono_imp1 out=impdata_reg_imp1 nimpute=1 seed=234;
     by _Imputation_;
     var base _W1;
     monotone reg(_W1=base);
run;
where _W1 represents the scheduled post-baseline scores Week 1
      BASE is the baseline score.
```

Since subjects from the active groups with non-missing values at time-point t are not included in the input dataset, they will not contribute to the estimation of an imputation model for time-point t. The imputation model will be estimated using control subjects only, while this call to PROC MI will impute missing data at time-point t for all subjects who need imputation at that time-point. This way, subjects from active group will be imputed based on the control subjects' model.

c) Then, we will assemble a dataset containing all subjects while including the just-imputed visit t data so that they can serve as predictors for the imputation of the next visit.

```
data impdata_mono1;
     set impdata_mono_rest1 impdata_reg_imp1;
run;
```

d) The a), b) and c) steps will be repeated for all others timepoints sequentially up to Week 12) using a reconstructed dataset (e.g., impdata_mono1) as a starting point in step (a).

Once all missing data will be imputed, then the dataset will be transposed and the change from baseline will be recalculated. Then, a MMRM model will be done by using PROC MIXED with by "_Imputation_" and then, the results will be combining by using PROC MIANALYZE.

A MMRM model will be applied by _Imputation_ as per model below:

```
proc sort data=impfinal; by _imputation_ avisitn; run;
ods output Lsmeans=lsmeans_est
            Diffs=diff
            FitStatistics=FitStat
            ConvergenceStatus=conv;
proc mixed data=impfinal_ order=internal;
  by _Imputation_;
  class usubjid trt01pn avisitn;
  model chg = trt01pn avisitn base trt01pn*avisitn base*avisitn / s ddfm=kr;
  repeated avisitn / subject = usubjid type = un;  /*type = cs*/
  lsmeans trt01pn*avisitn / cl pdiff;
run;
where USUBJID represents the subject identifier
      TRT01PN represents the planned treatment code (1: IGE 300 mg, 2: IGE 150 mg,
3: Placebo)
      AVISITN represents the study week from Week 1 up to Week 12
      BASE represents the baseline score
      CHG represents the change from baseline.
```

Note that the diff dataset created will include all the treatment interactions for each week at each imputation and lsmeans_est dataset created will include all LS Mean for each week and each treatment at each imputation.

Then, all results will be pooled by using PROC MIANALYZE based on diff and lsmeans_est datasets:

```
proc sort data=diff; by avisitn trt01pn _trt01pn; run;
ods output ParameterEstimates= diffEstimates;
proc mianalyze data=diff;
   by avisitn trt01pn _trt01pn;
   modeleffects estimate;
   stderr stderr;
   where avisitn=12 and _avisitn=12;
run;
```

diffEstimates dataset created will contain all statistics related to the comparison of LS Mean.

```
proc sort data=lsmeans_est; by avisitn trt01pn; run;
ods output ParameterEstimates= LSmeansEstimates;
proc mianalyze data=lsmeans_est;
   by avisitn trt01pn;
   modeleffects estimate;
   stderr stderr;
run;
```

LSmeansEstimates dataset created will contain all LS Mean for each treatment and each visit.

5.6.2.3 Logistic regression model

Binary endpoints will be estimated using logistic regression model using SAS procedure GLIMMIX, specifying the distribution as binary and link function as logit. Compute least square means and confidence intervals by treatment group. Also odds ratios and their confidence intervals are estimated between active treatment group and placebo group.

The following code will be used for this analysis:

```
ods output Estimates=diffor;
proc glimmix data=dataset order=internal;
   class treatment;
   model <response> (ref=first)= treatment baseline / dist=binary link=logit s ddfm=kr;
   lsmeans treatment / cl;
   estimate "150 mg - Placebo" treatment 0 1 -1 / cl exp;
   estimate "300 mg - Placebo" treatment 1 0 -1 / cl exp;
run;

where response = response analysis value (1=Yes/0=No)
      baseline = baseline score of the related parameter
      treatment = treatment group assigned in numeric (1 = 300 mg, 2 = 150 mg, 3 =
Placebo)
```

The odds ratios and their 95% CIs are given by the exponentiated variables in the diffor dataset (i.e. ExpEstimate, ExpLower, ExpUpper).

If the separation (at least one parameter estimate diverges to infinite, typically caused by zero or nearly zero response in the dichotomous covariates) happens, Firth's method (Heinze and

Schemper, 2002) implemented in PROC LOGISTIC will be used with option `firth clodds=pl` specified.

```
ods output Estimates=diffor;
proc logistic data=dataset order=internal;
  class treatment / param = glm;
  model <response> (ref=first)= baseline treatment / firth clodds=pl;
  estimate "150 mg - Placebo" treatment 0 1 -1 / cl exp;
  estimate "300 mg - Placebo" treatment 1 0 -1 / cl exp;
run;
```

5.6.2.5 Cox regression model

Time to event endpoints are analyzed using Cox regression model using SAS procedure PROC PHREG. Hazard ratios between active treatment group and placebo group and their confidence intervals will be generated by exponentially transforming the estimates and confidence intervals generated by the model.

The following code will be used for this analysis:

```
ods output Estimates=Est;
proc phreg data=dataset;
  class treatment / order=internal param=glm;
  model time*censor(1)= treatment baseline;
  lsmeans treatment / cl diff;
  estimate "150 mg - Placebo" treatment 0 1 -1 / cl exp;
  estimate "300 mg - Placebo" treatment 1 0 -1 / cl exp;
run;

where time = time to event (e.g. time to first ISS7 MID response by week 12)
      baseline = baseline score (e.g. of ISS7)
      censor = censoring indicator variable (1 = censored, 0 = event)
```

```
      treatment = treatment group assigned in numeric (1 = 300 mg, 2 = 150 mg, 3 =
Placebo)
```

The hazard ratios and their 95% CIs are given by the exponentiated variables in the Est dataset (i.e. ExpEstimate, LowerExp, UpperExp).

5.6.2.6 Multiple testing strategy

The results for multiple testing strategy will be presented as the example table below:

Hypo-thesis	Endpoint	Comparison	Estimate	95% CI	p-value		Outcome
					unadjusted	adjusted	
H1	ISS7 at Week 12	300 mg vs. Placebo	x.xx	(x.xx, x.xx)	0.007	0.014	Rejected
H2	ISS7 at Week 12	150 mg vs. Placebo	x.xx	(x.xx, x.xx)	0.049	0.098	Not rejected
H3	UAS7 at Week 12	300 mg vs. Placebo	x.xx	(x.xx, x.xx)	0.024	0.048	Rejected
H4~H14							
H15	Time to ISS7 MID response by Week 12	300 mg vs. Placebo	x.xx	(x.xx, x.xx)	0.023	0.170	Not rejected
H16	Time to ISS7 MID response by Week 12	150 mg vs. Placebo	x.xx	(x.xx, x.xx)	0.010	0.230	Not rejected

The following SAS macro will be used to derive the adjusted p-values and testing outcomes:

```
START gmcp(p, w, g);
    n = NCOL(p);  padj = j(1,n,0);  pmax = 0;  crit = 0;
    DO UNTIL(crit = n);
      pos = LOC(w>0);  zero = LOC(w<=0);  q = j(1,n,0);
      IF NCOL(pos)>0 THEN DO;
        q[pos] = p[pos]/w[pos];
        IF NCOL(zero)>0 THEN q[zero] = MAX(q[pos])+1;
      END;
```

```
      ELSE RETURN('gMCP ERROR: Disconnected hypothesis with weight 0.');
      rej = MIN(LOC(q - MIN(q) <= 0));
      padj[rej] = MAX(q[rej], pmax);
      pmax = padj[rej];
      g1 = J(n, n, 0);
      DO i = 1 TO n;
        w[i] = w[i] + w[rej]*g[rej,i];
        IF (g[i,rej]*g[rej,i]<1) THEN DO j = 1 TO n;
          g1[i,j] = (g[i,j] + g[i,rej]*g[rej,j])/(1 - g[i,rej]*g[rej,i]);
        END;
        g1[i,i] = 0;
      END;
      g = g1;  g[rej,] = 0;  g[,rej] = 0;  w[rej] = 0;  crit = crit + 1;
    END;
    padj = padj >< j(1,n,1);
    RETURN(padj);
FINISH;
```

The above macro used is published in Bretz et al (2011) and Alosh et al (2014), and is now a validated SAS macro that has been implemented in Novartis trials.

Inputs for *n* hypotheses:

- *p*: Observed *p* values (1 x *n* vector) with *n* equals to the number of hypotheses which is 16 here.
- *w*: Weights for the hypotheses (1 x *n* vector)

$$= (0.5, 0.5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$$

- *g*: Transition weights for the directed edges (*n* x *n* matrix):

$$= \begin{pmatrix} 0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1 \\ 0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \end{pmatrix}$$

The input *p* will contain two-sided unadjusted p-values, however if there are any comparisons that favor placebo, then the values of *p* will be adjusted in the following way before using the macro:

- For the hypotheses on the left hand side of the graphical approach (H1, H3, ..., H15, for the 300mg dose), find the first hypothesis where the direction is in favor of placebo (no matter what the p value is), then in the vector *p* replace the unadjusted p-values of all the following hypotheses with 1.
- For the hypotheses on the right hand side of the graphical approach (H2, H4, ..., H16, for the 150mg dose), find the first hypothesis where the direction is in favor of placebo (no matter what the p value is), then in the vector *p* replace the unadjusted p-values of all the following hypotheses with 1.

This is because the significance level of a rejected hypothesis is only passed on according to the graphical procedure for the test of another hypothesis if the treatment effect is in favor of omalizumab (see section 2.6.2).

Outputs:

- *padj*: Adjusted *p* values (1 x *n* vector)

The source program is located in GPS (Global Programming and Statistics).

5.7 Rule of exclusion criteria of analysis sets

Subject classification will be done by Protocol Deviation and non-Protocol Deviation criteria based on the tables below:

Table 5-2 List of protocol deviations that cause subjects to be excluded

Deviation ID	Description of Deviation	Excluded from
INCL01A / INCL01B	Informed consent missing or not signed prior to initiating study procedures.	FAS, SAF, PPS
INCL02	Age criteria not met	PPS
INCL03A	No diagnosis of CSU refractory to H1AH due to date of diagnosis not meeting criteria	PPS
INCL03B	No diagnosis of CSU refractory to H1AH due to lack of criteria on itch and hives	PPS
INCL03C	No diagnosis of CSU refractory to H1AH due to UAS7 score	PPS
INCL03D	No diagnosis of CSU refractory to H1AH due to UAS score	PPS
INCL03E	Criteria requirement at screening for prior use of H1AH at approved dose not met	PPS
INCL04	Unwilling or unable to complete a daily symptom eDiary	PPS
INCL05	Missing one or more eDiary entries in the 7 days prior to randomization	PPS
EXCL01	Underlying etiology for chronic urticarias other than CSU	PPS
EXCL03	Any other skin diseases than CSU with chronic itching	PPS
EXCL04	Previous treatment with omalizumab	PPS
EXCL05	Contraindications to diphenhydramine	PPS
EXCL06	History of anaphylactic shock	PPS
EXCL08	Patients who are sucrose intolerant	PPS
EXCL11	Inability to comply with study and follow-up procedures	PPS
EXCL13	Patients previously randomized into this study	FAS, PPS
EXCL14	Use of other investigational drugs	PPS
EXCL15	History of hypersensitivity to study drug or similar chemical classes	PPS

Deviation ID	Description of Deviation	Excluded from
EXCL17	Ongoing use of prohibited treatments at screening epoch	PPS
COMD01	Use of prohibited medication during randomized-treatment epoch.	PPS
TRT01	Incorrect dose of study medication	PPS
TRT02	Subject received the wrong study medication.	PPS
TRT03	Subject missed the administration	PPS
OTH01	Subject added new H1AH to the stable H1AH regimen during screening or randomized-treatment epoch.	PPS
OTH02	Subject decreased H1AH regimen during screening and treatment epoch or discontinued stable H1AH regimen on any epoch	PPS
OTH02A	Subject received overdose of diphenhydramine rescue medication during screening or randomized-treatment epoch.	PPS
OTH06	Patient not compliant with eDiary completion requirements	PPS
OTH07	Inappropriate maintenance of study medication	PPS
OTH09	Incorrect background treatment use of sedating H1AH	PPS
OTH10	Rescue medication mistakenly used as background medication	PPS
OTH11	Patient was Randomized but no study drug was taken	FAS, SAF, PPS

Table 5-3 Non-Protocol deviations that cause subjects to be excluded

Non-Protocol deviation criteria that cause a subject to be excluded	Excluded from
No treatment taken	FAS, SAF, PPS

Less than 3 doses of study drugs by Day 85	PPS

5.8 Systemic treatments

The list of the systemic treatment to use is:

- Systemic corticosteroids
- Hydroxychloroquine
- Immunosuppression (e.g. methotrexate, cyclosporine, tacrolimus, mycophenolate mofetil, cyclophosphamide, or triperygium wilfordii Hook)
- Intravenous immunoglobulin G
- H2 antihistamine
- Leukotriene receptor antagonist (LTRA)
- Oral Chinese traditional medicine prescribed for CSU

5.9 Latent TB status at baseline

An incubation issue has been identified which impacted the TB test samples in around 150 patients. The issue was that TB samples incubator temperature was fluctuating. EXCL21C Protocol Deviation (Potential invalid QuantiFERON TB negative test result) has been created in order to identify these subjects.

In addition, several sites have not specified the correct TB status at the time of randomization, they wrote “Yes” while the correct TB status was “No”. For all analysis performed, the latent TB status at baseline from laboratory data will be used (rather than IRT data).

6 Reference

Alosh, M., Bretz, F., & Huque, M. (2014). Advanced multiplicity adjustment methods in clinical trials. Statistics in medicine, 33(4), 693-713.

Bretz F, Maurer W, Brannath W et al. (2009) A graphical approach to sequentially rejective multiple test procedures. Stat Med; 284):586-604.

Bretz, F., Maurer, W., & Hommel, G. (2011). Test and power considerations for multiple endpoint analyses using sequentially rejective graphical procedures. Statistics in medicine, 30(13), 1489-1501.

Heinze, G. and Schemper, M. (2002) A Solution to the Problem of Separation in Logistic Regression, Statistics in Medicine, 21, 2409–2419.

Mallinckrodt, CH, Scott CW, and David SR (2001) Accounting for dropout bias using mixed-effects models. Journal of Biopharmaceutical Statistics; 11, 9-21.

Ratitch, B. and O'Kelly, M. (2011), Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures, in Proceedings of PharmaSUG 2011 (Pharmaceutical Industry SAS Users Group), SP04, Nashville.

Rubin DB (1976) Inference and Missing Data. Biometrika; 63:581-92.