Statistical Analysis Plan

Study Title:	A Randomized, Double-Blind, Placebo-Controlled Phase II Study to Evaluate the Efficacy and Safety of SHR0302 in Adult Patients with Moderate to Severe Atopic Dermatitis

Protocol No.: RSJ10303

Compound No.: SHR0302

NCT number: NCT04162899

Identification/Version No.: Final 1.0

Effective Date: 24Sep2020

Description:

The purpose of this Statistical Analysis Plan (SAP) is to describe the analyses planned for Protocol RSJ10303 and the information to be provided in the Statistical Analysis Report (SAR) and Clinical Study Report (CSR).

This SAP will elaborate the statistical analysis planned for the study from the aspects of efficacy and safety of multiple doses of the drug based on the study objectives. The study team members will have access to this document to understand what is included in the statistical analysis of the deliverables.

Topics:

SHR0302, Adult Patients with Moderate to Severe Atopic Dermatitis, Efficacy and Safety

Name and Affiliation of Author(s):

[redacted]	Senior Principal Statistician [redacted]
[redacted]	Statistician [redacted]

Signature Page for Review and Approval of Statistical Analysis Plan

Biostatistics, [redacted]		
Name and Title	**Signature**	**Date (mm-dd-yyyy)**
[redacted] Executive Director, Biostatistics		

Sponsor: Reistone Biopharma Company Limited		
Name and Title	**Signature**	**Date (mm-dd-yyyy)**
[redacted] Director, Biostatistics		

TABLE OF CONTENTS

Table 1. Abbreviations

Abbreviation	Definition
AD	Atopic dermatitis
AUC	Area under the curve
BMI	Body mass index
B-hCG	β-human chorionic gonadotropin
CI	Confidence interval
Cmax	Maximum plasma concentration
CRF	Case Report Form
CRP	C-reactive protein
DLQI	Dermatology life quality index
EW	Early withdrawal
ECG	Electrocardiogram
EMA	European Medicines Agency
EASI	Eczema area and severity index
ECRF	Electronic Case Report Form
FAS	Full analysis set
FDA	U.S. Food and Drug Administration
FSH	Follicle stimulating hormone
FVC	Forced vital capacity
GCP	Good Clinical Practice
GFR	Glomerular filtration rate
IB	Investigator's Brochure
IL	Interleukin
IGA	Investigator’s global assessment
JAK	Janus kinase
IEC	Independent Ethics Committee
IFN-γ	Interferon-γ
IgE	Immunoglobulin E
IgM	Immunoglobulin M
IND	Investigational new drug
IRB	Institutional Review Board
ITT	Intention-to-treat
IVRS	Interactive voice response system
IUD	Intrauterine device
LDH	Lactate dehydrogenase
LLoQ	Lower limit of quantitation
Mcg	Microgram
MedDRA	Medical dictionary for regulatory activities
NRS	Pruritus numerical rating scale
PK	Pharmacokinetics

Abbreviation	Definition
PP	Per-Protocol
PPD	Purified protein derivative
PT	Preferred term
QD	Once daily
RA	Rheumatoid arthritis
RAP	Reporting and Analysis Plan
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SAR	Statistical Analysis Report
SD	Standard deviation
SOC	System organ class
SPM	Study Procedures Manual
SCORAD	Atopic dermatitis score
SS	Safety set
TCS	Topical corticosteroids
TSH	Thyroid stimulating hormone
ULN	Upper limit of normal

Trademark Information

Trademark(s) of Reistone Biopharma Company Limited

Trademarks that do not belong to Reistone Biopharma Company Limited
SAS
MedDRA
WHODrug

1. INTRODUCTION

The purpose of this Statistical Analysis Plan (SAP) is to describe the analyses to be included in the Clinical Study Report for Protocol RSJ10303 V2.0 (Dec. 31, 2019).

All decisions regarding the final analysis (as defined in this SAP) will be made before all subjects have completed study drug administration and follow-up, and database freeze. In case of any change in the analysis described in this SAP, such change will be recorded in the Clinical Study Report together with the reasons.

2. DISCREPANCIES BETWEEN STATISTICAL ANALYSIS PLAN AND PROTOCOL

The changed contents in SAP relative to the analysis contents specified in the protocol and reasons are shown in Table 2.

Table 2. Discrepancies between Statistical Analysis Plan and Protocol

Protocol	Statistical Analysis Plan	Reasons to Change
Prior medications are defined as medications that were stopped prior to first study treatment.	Prior medications are defined as drugs used before the first study treatment.	The drugs used before the first administration of study drug until after the first administration of study drug will be classified into the prior and concomitant medications, both.
	The analyses of secondary efficacy endpoint, change from baseline and related analyses about absolute eosinophil counts are added.	The absolute eosinophil counts could be used as a secondary endpoint reflecting the efficacy.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1 STUDY OBJECTIVES

3.1.1 Primary objective

- To evaluate the efficacy of oral administration of SHR0302 at doses of 4 mg and 8 mg, QD vs. Placebo in adult subjects with moderate to severe atopic dermatitis at week 12.

3.1.2 Secondary objectives

- To evaluate the safety and tolerability of oral SHR0302 in subjects with moderate to severe atopic dermatitis.
- To evaluate the impact of oral SHR0302 on other efficacy endpoints in subjects with moderate to severe atopic dermatitis and on the reported outcomes of subjects.

- To investigate the pharmacokinetic (PK) characteristics of SHR0302 in subjects with moderate to severe atopic dermatitis, and to explore the exposure-effect relationship.

3.2 STUDY ENDPOINTS

3.2.1 Primary endpoint

- The percentage of subjects who achieve IGA response at week 12. IGA response is defined as IGA score of 0/1 (complete or almost complete clearance of skin lesions) with an improvement in IGA score by ≥ 2 from baseline.

3.2.2 Secondary endpoints

3.2.2.1 Safety endpoints

- Laboratory parameters will be used to evaluate safety and tolerability.
- Incidences of adverse events (AEs)/serious adverse events (SAEs) will be used to evaluate safety and tolerability.
- Measurements of vital signs (blood pressure, heart rate, body temperature, and respiratory rate).
- Assessment of total lipids, including triglycerides, low density lipoproteins (LDL), and high-density lipoproteins (HDL).
- Assessment of thyroid functions: TSH, free T4, and free T3.
- Assessment of liver and kidney functions.
- 12-Lead ECG.

3.2.2.2 Efficacy endpoints

- The percentage change in EASI score from baseline at weeks 1, 4, 8, and 12.
- The percentage of subjects who achieve IGA response at weeks 1, 4, and 8.
- The percentage of subjects whose EASI scores improve by ≥ 50%, 75%, and 90% (EASI50, EASI75, EASI90) from baseline at weeks 1, 4, 8, and 12.
- Changes in scoring atopic dermatitis (SCORAD) from baseline at weeks 1, 4, 8, and 12.
- The percentage of subjects whose SCORAD improve by ≥ 50%, 75%, and 90% (SCORAD50, SCORAD75, SCORAD90) from baseline at weeks 1, 4, 8, and 12.
- The percentage change in Pruritus Numerical Rating Scale (NRS) from baseline at weeks 1, 4, 8, and 12.
- The percentage of subjects whose NRS improve by ≥ 3 points from baseline (NRS-3) at weeks 1, 4, 8, and 12.
- Changes in Dermatology Life Quality Index (DLQI) from baseline at weeks 1, 4, 8, and 12.
- Changes in biomarker immunoglobulin E (IgE) from baseline at weeks 1, 4, 8, and 12.
- Changes in absolute eosinophil count from baseline at weeks 1, 4, 8, and 12.

3.2.2.3 Pharmacokinetics endpoints

- Systemic exposure (i.e., concentration and area under the curve) of SHR0302 at steady state in subjects with atopic dermatitis.

4. STUDY DESIGN

4.1 OVERALL STUDY DESIGN

This 12-week, randomized, double-blind, placebo-controlled, 3-arm parallel, multi-center clinical study will enroll adult patients with moderate to severe atopic dermatitis. This study includes a 4-week screening period, followed by a 12-week blind treatment period, and a 2-week follow-up period after the last dose of the study drug. Primary endpoint will be evaluated at the end of the 12-week blind treatment.

At randomization, subjects who meet the study criteria will be randomly assigned to one of three study groups on a 1:1:1 basis: 2 active drug dose groups (SHR0302 4 mg QD and SHR0302 8 mg QD) and 1 placebo group, to receive blind treatment for 12 weeks. A total of 105 subjects will undergo randomization.

All subjects who complete the treatment period will undergo a 2-week follow-up period until week 14 (last visit). In the event of early withdrawal or discontinuation, subjects will also undergo a 2-week follow-up period.

A total of 7 outpatient visits will be conducted during this study as detailed in Figure 1.

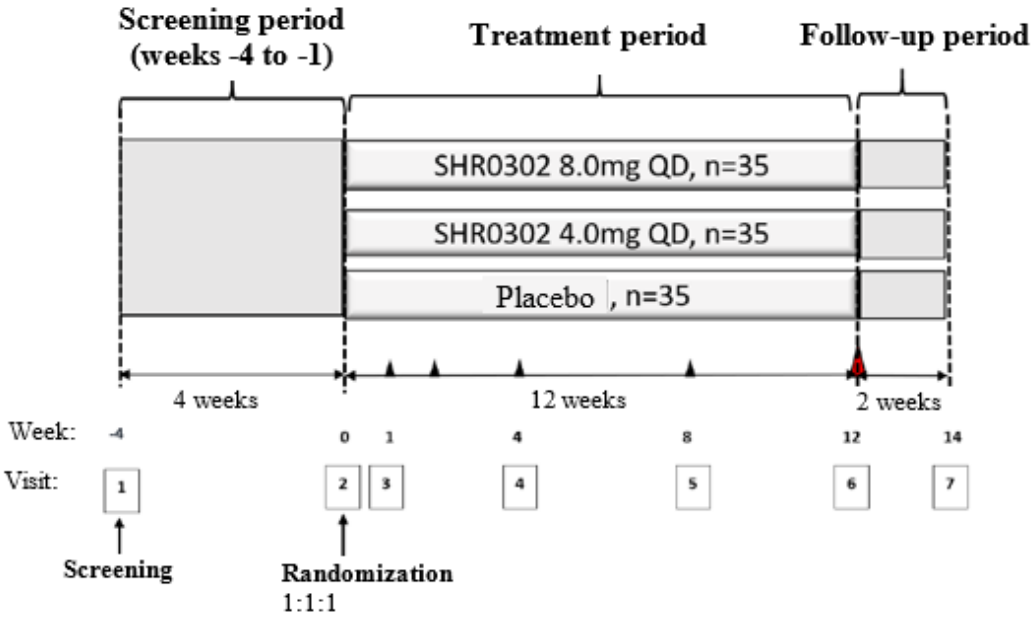

Figure 1. Study design diagram

4.2 DOSAGE REGIMEN

The study drug will be taken orally by subjects themselves in the morning, QD.

However, the outpatient dosing will be performed at the time of the first dose at baseline (Day 0) and at visits at weeks 4, 8, and 12. The study drug may be taken with or without food.

4.3 RANDOMIZATION AND BLINDING

This proposed study is a randomized, double-blind, placebo-controlled, 3-arm parallel, multi-center phase 2 study. At the randomized visit (Visit 2), if the subjects are in line with the inclusion criteria/not in line with any exclusion criteria, they would be randomized to the SHR0302 4 mg QD group, the SHR0302 8 mg QD group, or the placebo group at a ratio of 1:1:1. A computer-generated randomization table will be used to assign subjects to each study group via an electronic randomization system at the baseline visit. Subjects will be assigned with the subject numbers based on the order in which subjects are enrolled to participate in the study and will be told to keep their ID numbers throughout the study.

The double-blind method will be adopted regarding the study drugs taken during the 12-week treatment period. Neither subjects nor the study doctor will be aware of the study drugs taken by subjects.

The packages of the study drugs have been carefully considered and designed, so as to maintain the blindness and ensure the compliance of the subjects throughout the study. In all groups, the study drugs will be taken QD, and the following arrangements will be implemented for the dosage regimen and placebo (Table 3).

Table 3. Packaging design of study drugs in the treatment group

Study group	Dose from Day 0 to 84 (QD)
4 mg SHR0302	one 4 mg tablet + one placebo tablet
8 mg SHR0302	one 4 mg tablet + one 4 mg tablet
Placebo	one placebo tablet + one placebo tablet

At each dose time point (morning), subjects will swallow 2 tablets.

The investigator or the attending doctor could unblind the treatment assignments of the subject only in case of emergency (i.e., when it is necessary to know the study treatment in order to implement appropriate clinical treatment or for the sake of the health of subjects).

After the clinical database is locked, the statistician responsible for this project will complete the "Application Form for Randomization Code Release". All subjects will be unblinded once after the approval by the sponsor. The treatment allocation of unblinded subjects will be submitted to the SAS programmer for import into the analysis database for analysis.

5. PLANNED ANALYSES

5.1 INTERIM ANALYSIS

There is no formal interim analysis plan for this study.

5.2 FINAL ANALYSIS

The PK statistical analysis plan will not be included here as it will be described in a separate document.

The final analysis specified in this SAP will be performed after the database locked and the data released.

6. SAMPLE SIZE CONSIDERATIONS

This study is planned to enroll 105 subjects, 35 in each group, and provide approximately 82% power to test differences in primary endpoint between at least one active group and the placebo group, assuming the response rates in the 4 mg QD group, the 8 mg QD group and the placebo group are 26%, 26%, and 6%, respectively. By using the Hochberg incremental procedure, the type I error for the primary endpoint is controlled at the level of 0.1 bilaterally.

7. STATISTICAL ANALYSIS SET

The analysis sets for this study include full analysis set (FAS), per protocol analysis set (PPS), Efficacy sensitivity set (ESS) and safety set (SS).

- FAS: Includes all subjects who are randomized and received at least one time of the study drug administration. In FAS, subjects will be analyzed according to randomization, regardless of the actual treatment received.
- PPS: PPS is a subset of FAS, including subjects who have not experienced the major protocol deviations that significantly affect the efficacy evaluation, and have obtained primary efficacy evaluation endpoints (excluding remote visit results). Subjects excluded from the PPS will be identified and recorded before the study unblinding.
- ESS: ESS is a subset of FAS, including subjects who have not experienced the major protocol deviations that significantly affect the efficacy evaluation, and excluding subjects who have obtained the primary efficacy endpoints through remote visits. Subjects excluded from the ESS will be identified and recorded before the study unblinding.
- SS: Included all subjects who have received at least one dose of the study drug. The analysis set will be analyzed according to the actual treatment groups.

SS will be used for demographic and baseline characteristics analysis; FAS, PPS and ESS will be used for efficacy analysis, with FAS as the main analysis set and PPS and ESS as the supportive analysis sets; SS will be used for safety analysis. FAS, PPS, and ESS will be analyzed based on randomization groups, and SS will be analyzed based on the actual treatment groups.

Prior to treatment group unblinding and database release, data from all subjects will be assessed to determine the classification of each statistical analysis set.

8. STATISTICAL HYPOTHESIS AND MULTIPLICITY COMPARISON

In this study, there are multiple comparisons in the primary analysis. The hypothesis test will be performed for the primary efficacy endpoint, i.e., the comparison of 4 mg QD and 8 mg QD groups with placebo, respectively. The type I error of the primary endpoint will be controlled at the level

of 0.1 bilaterally by using the Hochberg incremental procedure[1]. The statistical hypotheses are detailed in the efficacy analysis section.

Unless otherwise specified, the significance level for other analyses will also be set to 0.1 bilaterally.

9. DESCRIPTION OF TREATMENT GROUPS

Unless otherwise specified, the statistical description and summary will be presented according to below treatment groups as follows.

Table 4 Description of Treatment Groups

Treatment Description	Final Data Display
4 mg tablet + placebo tablet	4 mg QD SHR0302
4 mg tablet + 4 mg tablet	8 mg QD SHR0302
Active drug total	Treatment Total
Placebo tablets + placebo tablets	Placebo
Total	Total

10. CONSIDERATIONS FOR DATA ANALYSIS AND DATA DEALING CONVENTIONS

10.1 GENERAL RULES FOR STATISTICAL DESCRIPTION

Unless otherwise specified, descriptive statistics appropriate to the type of variable will be provided.

Continuous variables will be presented with the number of non-missing values (n), mean, standard deviation (SD), median, minimum, and maximum values. If less than 3 observations, only n, minimum and maximum values will be presented.

Categorical variables will be presented as frequencies and percentages.

10.2 STATISTICAL RESULT REPORTING RULES

Analyses will be performed using SAS® version 9.4 or higher.

All analysis outcomes will be presented by statistical tables and/or figures. And the original data listings corresponding to the tables or figures will be provided, including important derivation variables and study endpoints, and unscheduled visit data. The general rule for sorting variable in listings is by study site, dose group/treatment group, subject ID, demographic characteristics, visit/treatment cycle, date and time, test/assessment, original results (unit), standard/derived results or study endpoints, explanation for results, reason and other information. The variable sort could be adjusted according to original data, data characteristics or requirement from report. Please refer to the mock shells for each listing, as not all of which will be specified in this analysis plan.

The decimal places are specified as below.

Table 5 Basic Descriptive Statistics

Name	Illustration	Decimal Places（dp）
N	Number of subjects in analysis set	Always present 0 dp
n	Number of non-missing observations	Always present 0 dp
%	Percentage	Categorical data 1 dp

Mean	Arithmetic Mean	1 more dp than original data
SD	Standard Deviation	2 more dp than original data
Median	Median	1 more dp than original data
Min.	Minimum value	Same with original data
Max.	Maximum value	Same with original data

Table 6 Derived Data Or Statistics of Estimated Parameters

名称	说明	小数位数（dp）
SE	Standard Error	1 more dp than derived data or parameter
95% CI	95% Confidence Interval	1 more dp than derived data or parameter
p value	P value	Present 3 dp, or < 0.001

The derived data and its statistics will be 1 more decimal place than the original data or the corresponding statistics.

Reasonable decimal places will be adjusted for proper presentation and interpretation.

11.DERIVED AND TRANSFORMED DATA

11.1 BASELINE DEFINITION AND CHANGE FROM BASELINE

Unless otherwise specified, the baseline value for a parameter of interest is the latest non-missing value prior to date of the first administration of study drug.

Change from baseline and percentage from baseline will be derived as follows:

- Change from baseline = the post-baseline value - the baseline value of a parameter of interest for a subject.
- Percentage change from baseline = (post-baseline value - the baseline value)/ baseline value of a parameter of interest for a subject.

If either the baseline or the post-baseline value is missing, then the change or percentage change will be set to missing.

11.2 STUDY DAY

Day 1 = Date of the first administration of study drug for a subject

Study Day

= Calendar Date – date of the first administration of study drug + 1 if the calendar date is on or after the date of the first administration of study drug, or

= Calendar Date - date of the first administration of study drug if the calendar date is before the first administration of study drug

11.3 MULTIPLE MEASUREMENTS AT THE SAME TIMEPOINT

If there are multiple measurement results at a timepoint, unless otherwise specified, use the mean of these measurements to summarize and derive. If there is unscheduled measurement, consider the

scheduled data firstly, if missing at the scheduled visit, then impute with closest non-missing per-protocol data within the visit window to make analysis.

11.4 EARLY WITHDRAWA

All subjects who early withdraw/discontinue study drug administration will be documented, and the reasons will also be recorded in the final clinical study report. All available data will be included in each applicable analysis set for summarization and analysis. The definition of each analysis set is detailed in Section 7.

If the study is terminated early, all available data will be listed and the study team will review these data to evaluate which statistical analysis will be still applicable.

11.5 MISSING DATA

In general, missing data will not be included in descriptive statistics analysis.

Missing efficacy data will be handled as specified in section of efficacy analysis.

For safety data, if there is "< x", "<= x", "> x", or ">= x" value in the quantitative safety data, the "x" will be used for derivation or summary. But the final handling and analysis will be based on medical opinion.

If the dates used for derivation or calculation are not complete or missing, the following rules will be considered.

- If the onset date of an AE is missing or incomplete, it is assumed to have occurred during the study treatment phase, except if the stop date indicates that the event happened before the study treatment.
- If the phase of a concomitant medication used cannot be identified, due to missing or incomplete onset/end date, it will be assumed to have occurred during both screening and study treatment phase, except if partial dates indicate that the medication is not administrated in one of the study phases.
- When calculating the age of subject, if birth date is missing or not complete, the age collected in CRF will be used, else impute with the first date of the month or the first day (01January) of the year.

12.DATA SOURCE

Detailed sources of all types of data that will be analyzed in this SAP are described as below.

Table 7 Data source

Data type	Origin	Data format	Responsible Party
Subject Demographics, Baseline Characteristics, Efficacy, Safety	eCRF	The SAS dataset (. sas7dat)	[redacted]

13.STUDY POPULATION ANALYSIS

13.1 DISPOSITION OF SUBJECTS

In all screened subjects, the number and percentage of subjects screened, screen failure, screened successfully and randomized, screened successfully and but not randomized will be summarized as well as the reasons.

In all enrolled subjects, the number and percentage will be summarized of subjects enrolled, randomized, with at least one dose of assigned drug, with week 12 remote visit, with week 12 prolonged visit and those who have completed study or discontinue from treatment or study along with reasons. The flow chart of subject disposition will be provided in a separate file.

The number of subjects in different analysis sets by treatment group will be summarized in enrolled subjects.

13.2 PROTOCOL VIOLATION

Major protocol deviations (including informed consent deviations, deviation of inclusion/exclusion criteria, deviation of termination/withdrawal criteria, study medication deviation, use of prohibited medications, trial procedures/procedures, and any other deviation deemed likely to have a significant impact on study results) will be summarized.

All major protocol deviations will be discussed and identified at the data review meeting and the impact on each analysis set will be assessed.

13.3 DEMOGRAPHIC AND BASELINE CHARACTERISTICS

The following demographic and baseline characteristic will be summarized. Continuous variables will use statistics including mean, standard deviation and so on, Number of non-missing values (n) and corresponding percentage (%) will be presented for descriptive presentation of categorical data.

- Continuous variables: age, height, weight, body mass index (BMI)
- Categorical variables: gender, ethnic group, race

  Age = INT((date of informed consent - date of birth)/365.25);

  Body mass index (BMI) = weight (kg)/height (m^2).

Baseline characteristics includes tobacco use (summary of smoking status and frequency. Smoking status includes never smoke, previous smoke and current smoke), alcohol use (summary of use status and frequency. Use status includes never use, previous use and current use), virology test (hepatitis B surface antigen, hepatitis C antibody, human immunodeficiency virus antibody), tuberculosis infection test (summary of both tests types and results), and primary diagnosis of atopic dermatitis (disease duration), atopic dermatitis disease activity (Eczema Area and Severity Index (EASI), BSA, Investigator's Global Assessment (IGA) and Pruritus Numerical Rating Scale (NRS). Atopic dermatitis disease activity will be summarized using baseline results.

Duration of atopic dermatitis = (date of initial dose of IMP– date of diagnosis of atopic dermatitis + 1)/365.25 in years. When the date of primary diagnosis for atopic dermatitis recorded in the eCRF is incomplete or missing, the following rules will be applied for imputation:

- If day is missing then 15th of the current month will be imputed, and if month is missing then July of the current year will be imputed;
- If duration of atopic dermatitis is less than 1 year due to the imputation, then 1st day will be used for missing days, and January will be used for missing months;
- Missing year will not be imputed;
- Listings of missing dates will be presented as collected on the eCRF.

Baseline results for other safety endpoints will be summarized in safety analysis part. Other efficacy endpoints will be summarized in the efficacy analysis part.

13.4 MEDICAL HISTORY

Medical history and concomitant illness will be presented for the safety analysis set.

Medical history and concomitant illness will be coded using Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class (SOC) and preferred term (PT).

Medical history is defined as those conditions ended prior to the date of the first administration of study drug; concomitant illness is defined as those conditions started prior to and ended on or after the date of the first administration of study drug. If end date of one condition could not show earlier than date of the first administration of study drug then it will be classified into concomitant illness.

13.5 TREATMENT OF ATOPIC DERMATITIS

Treatment of atopic dermatitis includes atopic dermatitis medication treatment and non-drug treatment. The number of patients treated, and types of treatment will be summarized by treatment group. Prior and concomitant medication treatment/non-drug treatment will be summarized separately. Prior medication treatment is defined as medication treatment ended prior to the date of initial dose of IMP. Concomitant medication treatment is defined as medication treatment administered at least one time after date of initial dose of IMP. In the case where it is not possible to define a medication as prior, concomitant, this medication will be classified as concomitant. Medications stopped on the same day as the first study treatment will be considered as prior medication only. Medications started prior and ended on date of initial dose of IMP will be classified into both prior and concomitant medication. Prior and concomitant non-drug treatment are defined in same manner with medication treatment.

13.6 PRIOR AND CONCOMITANT MEDICATIONS/NON-DRUG TREATMENTS

Prior and concomitant medication and non-drug treatment data are captured in the CRF. Prior and concomitant medications will be coded by WHODD and summarized by anatomical therapeutic chemical (ATC) classification system and preferred name (PN). Prior and concomitant non-drug treatment will be coded by MedDRA and summarized by system organ class (SOC) and preferred term (PT) respectively.

The definition of prior and concomitant medications/non-drug treatments are the same as prior and concomitant medications/non-drug treatments for atopic dermatitis.

14.EFFICACY ANALYSIS

14.1 OVERVIEW OF EFFICACY ANALYSES

The following hypotheses will be tested based on the primary endpoint and objective:

- H_{01}: Percentage of subjects achieving an IGA response in the SHR0302 8 mg QD dose group at Week 12 - Percentage of subjects achieving an IGA response in the placebo group at Week12 = 0;
- H_{11}: Percentage of subjects achieving an IGA response at Week 12 in the SHR0302 8 mg QD dose group - Percentage of subjects achieving an IGA response at Week 12 in the placebo group $\neq$ 0.
- H_{02}: Percentage of subjects achieving an IGA response in the SHR0302 4 mg QD dose group at Week 12 - Percentage of subjects achieving an IGA response in the placebo group at Week12 = 0;
- H_{12}: Percentage of subjects achieving an IGA response at Week 12 in the SHR0302 4 mg QD dose group - Percentage of subjects achieving an IGA response at Week 12 in the placebo group $\neq$ 0.

By using the Hochberg[1] incremental procedure, the type I error for the primary endpoint will be controlled at the level of 0.1 bilaterally. Two p-values will be gained and sorted ($p_{(2)} \geq p_{(1)}$). Then start with the least significant comparison ($p_{(2)}$) and continue as long as the test is not significant until the first time when a significant comparison occurs, and all remaining hypotheses will be rejected. The decision rules are provided in below figure 2.

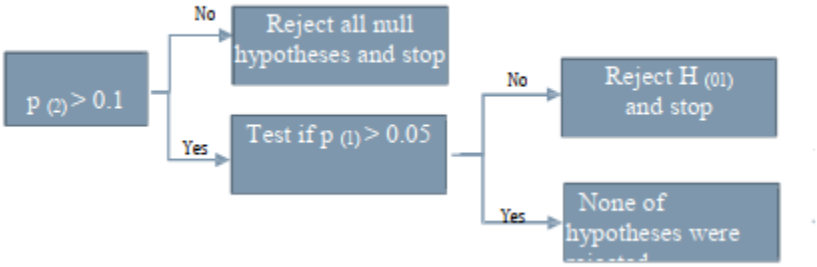

Figure 2 The decision rules of hypothesis

NOTE: $H_{(01)}$ is the corresponding null hypothesis of the smallest p value p(1)

Analyses of secondary efficacy and other analyses will be provided for comparison 4 mg QD and 8 mg QD versus placebo, respectively.

14.2 PRIMARY EFFICACY ANALYSIS

Analysis of the primary efficacy endpoint will be based on FAS.

The primary efficacy endpoint for this study is the percentage of subjects achieving IGA response at Week 12. IGA is the investigator's global assessment of the subject's general condition and consists of a 6-point severity scale from clear to very severe disease ("complete clearance", "almost

complete clearance", "mild", "moderate", "severe", and "very severe"). IGA response is defined as IGA score of 0/1 (complete or almost complete clearance of skin lesions) with an improvement in IGA score by ≥ 2 from baseline.

Subjects who early withdraw from the study for any reason will be considered as non-responders at each post-withdrawal assessment time point. Descriptive summaries of the primary efficacy endpoint will be based on non-missing data only.

Due to the impact of the COVID-2019 pandemic, subjects may have their primary efficacy endpoint assessed outside the scheduled time window of visit 6. If the primary efficacy endpoint is missing within the planned time window (between study days 82 and 88), it will be considered to impute with results of assessment within 89 to 102 study day (ie, extended visit).

The method of normal approximation to binomial will be used to test whether each active drug group is superior to the placebo group. Estimation of the difference and its 90% confidence interval between each drug group and the placebo group, as well as the corresponding p value for test will be provided.

Referenced SAS code:

PROC FREQ DATA = adam.adeff;

*TABLE treat*aval/CHISQ RISKDIFF (CL = WALD) ALPHA = 0.1;*

RUN;

In the above codes, variables are defined as follows:

Aval = IGA response at Week12

Treat = treatment group

In addition, IGA response and its 90% confidence interval at Week 12 will also be provided for each treatment group.

14.3 SECONDARY EFFICACY ANALYSES

The secondary efficacy endpoints in this study are defined in section 3.2.2.

Unless otherwise specified, analyses of secondary efficacy endpoints are based on the FAS.

For dichotomous secondary efficacy endpoints, subjects who early withdraw from the study for any reason will be considered as non-responders at each post-withdrawal assessment time point. Descriptive summaries of secondary efficacy endpoints will be based on non-missing data only.

For continuous secondary efficacy endpoints, missing values will not be imputed.

For comparison of dichotomous endpoints, if both the expected number of subjects achieving response and the expected number of subjects not achieving response are ≥ 5 in each treatment group, the normal approximation method will be used for analysis; otherwise, Fisher's exact method described in Section 14.4.1 will be used instead.

14.3.1 Percentage of subjects achieving IGA response at Week 1, Week 4, Week 8

Rates of IGA response at other specified efficacy assessment point (Week 1, Week 4, and Week 8) will analyzed using the method of normal approximation to binomial or Fisher's exact test.

14.3.2 Percentage of EASI change from baseline at Weeks 1, 4, 8, and 12

The EASI score is the sum of the skin lesion severity scores and the integral and coefficient product of the area affected by the skin lesions in the four parts of the body: head and neck, trunk, upper limbs, and lower limbs. The EASI score ranges from 0.0 to 72.0 and can be varied in increments of 0.1, Higher scores indicate greater severity.

A mixed-effects repeated measures model (MMRM) will be used to analysis percentage of EASI change from baseline including treatment group, baseline value, visit, and treatment-by-visit interaction as fixed effect and visit as repeated variable. An unstructured covariance matrix will be used to model the within-subject variance-covariance errors. If model fails to converge , the following options will be considered in order until convergence achieved : TOEPH (Heterogeneous Toeplitz) 、 AR(1)(Heterogeneous autoregressive) 、 TOEP(Toeplitz) and CS (Compound symmetry). The Kenward-Roger approximation will be used to estimate the degrees of freedom. The comparisons of interest will be the contrasts between each active dose and placebo at every time point. For the treatment difference of each active dose versus placebo at each visit, the model estimated least square means, standard error, p-value, and two-sided 90% confidence interval will be reported.

Referenced MMRM SAS code:

```
PROC MIXED DATA=adam.adeff METHOD=REML;
  CLASS treat visit subj;
  MODEL chgEASI= treat visit treat*visit basval /SOLUTION DDFM=KR;
  REPEATED visit/SUBJECT=subj TYPE=UN;
  LSMEANS treat*visit/DIFF PDIFF CL;
RUN;
```

In the above code, variables are defined as follows:

ChgEASI = change from baseline in EASI for each visit

Treat = treatment group

Visit = visit

Subj = subject

Basval = baseline EASI score

Least square means of change from baseline in EASI score at each assessment visit will be graphically displayed for each treatment using a line plot.

EASI score and change from baseline will be summarized by treatment at each efficacy assessment visit, and statistics including mean, standard deviation, and 90% confidence interval will be provided. Same statistics will also be provided for each site score as well as for the four symptoms.

14.3.3 Percentage of subjects with ≥ 50%, 75%, and 90% EASI improvement from baseline (EASI50, EASI75, EASI90) at Weeks 1, 4, 8, and 12

Percentage of EASI change from baseline will also be derived as a dichotomous result according to the following three criteria: ≥ 50% (EASI50), ≥ 75% (EASI75), and ≥ 90% (EASI90). For each derived dichotomous EASI result, the comparison between 4 mg QD and 8 mg QD versus placebo will be performed by method of normal approximation to binomial or Fisher's exact test, respectively, according to actual data distribution.

The number and percentage of subjects achieving EAS50, EAS75, and EAS90 will also be summarized for each treatment group at each efficacy assessment visit.

14.3.4 Change from baseline in SCORAD at Week 1, Week 4, Week 8, and Week 12

The SCORAD score integrates the observational assessment of physicians and patients. The scoring scale consists of three major sections: the objective signs section includes lesion area and lesion severity, and the subjective symptoms include the degree of itching and sleep effects. SCORAD scores range from 0 to 103.

SCORAD quantitative results will be analyzed using the same MMRM method as EASI.

The SCORAD scores and change from baseline at each assessment visit will be summarized, and the scores of each objective sign sections (lesion area, lesion severity) and the subjective symptom (pruritus, impact on sleep) will be summarized using statistics including mean, standard deviation and 90% confidence interval. A line graph will be used to graphically display the least squares mean of change from baseline in SCORAD score for each treatment assessment visit (including five scales: Total Score, A Lesion Area, B Lesion Severity, C Pruritus and sleeplessness).

14.3.5 Percentage of subjects With ⩾ 50%, 75%, and 90% SCORAD improvement from baseline (SCORAD50,SCORAD75, SCORAD90) at Weeks 1, 4, 8, and 12

Percentage of SCORAD change from baseline will also be derived as a dichotomous result based on the following three criteria: ≥ 50% (SCORAD50), ≥ 75% (SCORAD75), and ≥ 90% (SCORAD90). For each derived dichotomous SCORAD result, the comparison between 4 mg QD and 8 mg QD versus placebo will be performed using method of normal approximation to binomial or Fisher's exact test, respectively, according to actual data distribution.

The number and percentage of subjects achieving SCORAD50, SCORAD75 and SCORAD90 will also be reported.

14.3.6 Percent of pruritus numerical Rating scale (NRS) change from baseline at Weeks 1, 4, 8, and 12

At each visit, the degree of pruritus/scratching is assessed by the subject on an 11-grade rating scale from 0 (no pruritus/scratching) - 10 (severe pruritus/scratching) for the first 24 hour of the visit.

The NRS quantitative results will be analyzed using the same MMRM method as for EASI.

The NRS score and change from baseline at each efficacy assessment visit will be summarized by treatment and mean, standard deviation and 90% confidence interval will be provided.

14.3.7 Percentage of subjects achieving pruritus NRS improvement ⩾3 from baseline in at Weeks 1, 4, 8, and 12

Percentage of NRS change from baseline will also be derived into two categories: ≥ 3 points and < 3 points. Dichotomized NRS results will be compared between 4 mg QD and 8 mg QD versus placebo groups using method of normal approximation to binomial or Fisher's exact test, respectively, according to actual data distribution, noting that the comparison of NRS differences will only be performed in the population with baseline NRS score ≥ 3.

The number and percentage of NRS will also be reported as follows: 0, 1-3, 4-6, 7-9 and 10 points for 5 grades.

14.3.8 Change from baseline in dermatology life quality index (DLQI) at Weeks 1, 4, 8, and 12

The DLQI is a dermatology-related questionnaire that allows subjects to assess the impact of skin disease on quality of life through 10 questions with total score ranging from 0 to 30.

The rules for scoring the 10 questions are as follows:

- Not at all: 0 points (Item 7 Question: No (Not at all), 0 points);
- A little: 1 point (question 7: no (a little), 1 point);
- Many: 2 points (question 7: No (many), 2 points);
- Very much: 3 points (question 7: Yes, 3 points).

Change from baseline in DLQI will be analyzed using the same MMRM model as the EASI.

The DLQI score and change from baseline at each treatment visit will be summarized using statistics including mean, standard deviation, and 90% confidence interval.

14.3.9 Change from baseline in biomarker IgE at Weeks 1, 4, 8 and 12

Atopic dermatitis is divided into endogenous and exogenous types. Patients with exogenous AD usually present with elevated serum total IgE.

Change from baseline in IgE will be analyzed using the same MMRM model as the EASI.

IgE and change from baseline will be summarized by treatment group at each efficacy assessment visit using statistics including mean, standard deviation and 90% confidence interval.

14.3.10Change in absolute eosinophil count from baseline at Weeks 1, 4, 8, and 12

The change from baseline in absolute eosinophil count will be analyzed using the same MMRM model as described for the EASI.

Absolute eosinophil counts and changes from baseline will be summarized by treatment group at each efficacy assessment visit using statistics including mean, standard deviation, and 90% confidence intervals.

14.4 SENSITIVITY ANALYSES

14.4.1 Statistical analysis methods

The primary efficacy endpoint of IGA response at Week 12 specified in Section 14.2 will be tested based on the full analysis set (FAS) using Fisher's exact test together with Exact Unconditional 90% confidence interval of difference between active groups versus placebo group. The 90% confidence interval of IGA response rate at Week 12 based on Clopper-Pearson method will also be reported for each treatment group.

Referenced SAS code:

```
PROC FREQ DATA = adam.adeff;
TABLE treat*aval/CHISQ RISKDIFF ALPHA = 0.1;
  EXACT RISKDIFF (METHOD = SCORE);
RUN;
```

14.4.2 Analysis set

Sensitivity analyses of the primary efficacy endpoint (IGA response at Week 12), EASI50, EASI75, EASI90, and EASI quantitative results at Week 12 will be performed based on the per protocol set (PPS) and the efficacy sensitivity analysis set (ESS) using the same analysis methods described in Section 14.2, Section 14.3.3, and Section 14.3.2, respectively.

14.4.3 Remote assessment

Considering the influence of assessment on efficacy evaluation by remote visits (i.e., efficacy endpoints are assessed at non study sites), the analyses of the primary efficacy endpoint IGA response at Week 12, and the secondary endpoints of EASI50, EASI75, and EASI90 at Week 12 will also be evaluated based on the full analysis set (FAS) using the same analysis methods as described in section 14.2, section 14.3 according to different strategies handling remote visit such as treated as missing or imputed (see section 14.4).

14.4.4 Handling of missing data

Missing data will be considered as non-response as described in Section 14.2, Section 14.3, for dichotomous efficacy endpoints. The impact of different imputation methods dealing with missing data such as last observation carried forward (LOCF) will also be further explored based on the full analysis set (FAS) to verify robustness of conclusion.

Last observation carried forward (LOCF): If a patient has no observation at a post-baseline visit (missing), then it will be filled in by last valid observation. Note that only scheduled efficacy assessment results are included in LOCF imputation.

Analysis data sets chosen, rules of handling remote visit assessment and missing data for primary and sensitivity analyses of dichotomous efficacy endpoints are specified in table 8.

Table 8 Summary of analysis for dichotomous efficacy endpoints

	Remote evaluation			Missing data			Prolonged visit	Evaluation endpoint
	Used	Non-response	Deleted	Non-response	LOCF	Deleted	Used	
Primary analysis (FAS)	Y			Y			Y	IGA at week 12, (Method of normal approximation to binomial)
Secondary analysis (FAS)	Y			Y			Y	All dichotomous efficacy endpoints
Sensitivity analysis 1 (PPS)			Y			Y	Y	IGA at week 12, EASI50/75/90 at week 12
Sensitivity analysis 2 (FAS)	Y			Y			Y	IGA at week 12 (exact test)
Sensitivity analysis 3 (FAS)		Y		Y			Y	IGA at week 12, EASI50/75/90 at week 12
Sensitivity analysis 4 (FAS)	Y				Y		Y	IGA at week 12, EASI50/75/90 at week 12
Sensitivity analysis 5 (ESS)			Y	Y			Y	IGA at week 12, EASI50/75/90 at week 12

15.SAFETY ANALYSIS

Safety evaluation will be performed based on SS. Safety evaluation endpoints includes AEs, SAEs, laboratory test results (hematology, blood biochemistry, urinalysis, lipid test, and TSH), vital signs (including blood pressure, pulse, body temperature, and respiratory rate), electrocardiograms, and physical examinations. The duration of medication, total drug exposure information and medication compliance will be summarized.

All safety data will be summarized by treatment group and total.

15.1 STUDY DRUG EXPOSURE AND TREATMENT COMPLIANCE

The total days of actual dosing, total amount of actual dosing and total amount of planned dosing for study drug will be summarized by treatment group using statistics including means, standard

deviation and so on. Proportion of subjects with overdose and omission/interruption will also be provided. Treatment compliance will be calculated based on the administration plan and actual total amount and summarized using statistics including means, standard deviation and so on. The proportion of subjects with treatment compliance < 80%, 80% -120%, and > 120% will be reported, respectively.

The number and percentage of overdoses (i.e., actual daily dose greater than planned), overdoses leading to adverse events (i.e., actual daily dose greater than planned and resulting in clinical signs and symptoms) will also be summarized separately.

The total days of actual dosing, total amount of actual dosing, average amount of actual dosing for study drug and treatment compliance will be derived as follows:

- The total days of actual dosing (days) = (end date 1 - start date 1 + 1) + (end date 2 - start date 2 + 1) +... + (end date of n- start date n + 1), note that only records with actual daily tablets > 0 will be summarized;
- The total amount of actual dosing (tablets) = (end date 1 - start date 1 + 1) * actual daily tablets 1 + (end date 2 - start date 2 + 1) * actual daily tablets 2 +... + (end date n- start date n + 1) * actual daily tablets n;
- The total amount of planned dosing (tablets):
    - If early withdrawal occurs before the end date of the treatment period (i.e., withdrawal before Week 12), The total amount of planned dosing (tablets) = (date of early withdrawal - date of first dose + 1) *2;
    - If not withdrawn or withdrawn after the end date of the treatment period, The total amount of planned dosing (tablets) = min {max (date of last dose - date of first dose + 1, 82), 88} 2;
- Treatment compliance (%) = The total amount of actual dosing (tablets)/ The total amount of planned dosing (tablets).

15.2 ADVERSE EVENTS

According to the Medical Dictionary for Regulatory Activities (MedDRA) (Version 21.0 or later), AEs will be coded using the system organ class (SOC) and preferred terms (PT). Drug related AEs includes "definitely related", "probably related", and "possibly related" AEs. Missing relationship to study drug will be classified as drug-related AEs. The investigator will grade the severity of each adverse event as mild, moderate, and severe. Missing severity will be treated as moderate.

AE will be summarized by SOC and PT. The incidence rate of AE in summary table will be sorted by descending order among SOCs and within a SOC. And the incidence rate is based on the number of subjects with AE(s), not the number of events. If a subject happened more than one same AEs, in the SOC or PT of this AE, only count once for this subject.

AEs during screening phase are defined as any AEs with before the time of start of investigational product administration. Treatment-emergent adverse events (TEAEs) are defined as AEs which commence on or after the time of start of investigational product administration. In the cases where date of AEs is missing or incomplete and it is not possible to define an AE as treatment emergent

or not, the AEs will be classified TEAEs. Serious infections are defined as any infection (viral, bacterial, and fungal) requiring hospitalization or parenteral antimicrobials.

15.2.1 All adverse events

The following summaries will be provided for all adverse events (AEs):

- Summary of all AEs, screening AEs, TEAEs, including incidence and number of events;
- Summary of all TEAEs by Maximum severity;
- Summary of all study drug related TEAEs;
- Summary of all study drug related TEAEs by maximum severity.

15.2.2 Serious adverse event

The following summaries will be provided for all serious adverse events (SAEs):

- Summary of all treatment emergent SAEs;
- Summary of all study drug related treatment emergent SAEs.

15.2.3 Death

The following summaries will be provided for all deaths:

- Summary of TEAEs with an outcome of death.

15.2.4 Adverse events affecting administration of study drug or leading to study withdrawal

The following summaries will be provided:

- Summary of all TEAEs affecting administration of study drug (including dose increased, dose reduced, drug interrupted, and drug withdrawn);
- Summary of all TEAEs leading to withdrawal from study.

15.2.5 Adverse events of special interest

The following summaries will be provided for treatment-emergent adverse events of special interest:

- Serious infection
    - Summary of all treatment-emergent serious infections
    - Summary of all drug-related treatment-emergent serious infections
- Thyroid disorders (hypothyroidism, hyperthyroidism)
- Hepatic injury

15.3 CLINICAL LABORATORY EVALUATIONS

Laboratory tests include hematology, biochemistry, urinalysis, lipid profile test and thyroid function test.

The quantitative results and change from baseline will be summarized for each laboratory item at each evaluation time point. For categorical data, the number and percentage of subjects in each category will be provided.

Shift table from baseline according to normal range criteria will summarize categorical results of by visit.

Shift tables representing categorical change of laboratory results (normal and abnormal, according to normal range criteria) from baseline to each post will be provided. Shift table will also summarize categories of clinical significance from baseline to worst post-baseline value (abnormal with clinically significant > abnormal without clinically significant >normal>missing), and unscheduled visit will be included.

Line plot of the following laboratory tests over time will be provided together with its mean and 90% confidence interval by treatment group.

- Hematology: hemoglobin, absolute lymphocyte count, absolute neutrophil count, and platelet count;
- Lipid profile tests: total cholesterol, low density lipoprotein, high density lipoprotein, and triglycerides;
- Thyroid function tests: thyroid stimulating hormone (TSH), free T3 (fT3) and free T4 (fT4).

The following hepatic events will also be summarized by visit for liver chemistry:

- ALT or AST > 8 × ULN, or
- ALT or AST > 5 × ULN, or
- ALT or AST > 3 × ULN with elevated total bilirubin > 2 × ULN.

Listings of laboratory parameters will display observed values and change from baseline, if applicable, with the flag for values, which are outside of normal range. And the listings of abnormal values will be provided with flag for clinically significant abnormalities. listings containing only clinically significant laboratory values will also be provided. If there are no clinical significance evaluations, listings of abnormal values will be provided instead.

15.4 VITAL SIGNS

Vital signs to be summarized include blood pressure, heart rate, body temperature, and respiratory rate. Observed values and change from baseline will be summarized using descriptive statistics. A subject listing of all vital signs will also be presented.

15.5 ELECTROCARDIOGRAM (ECG)

ECG parameters to be analyzed include heart rate (HR), PR interval, QRS tachycardia, QT interval, and corrected QT interval (QTcF).

Triplicate ECG measurements are collected at each visit. The averaged value of quantitative 3 ECG results and worst clinical assessment (abnormal clinically significant > abnormal not clinically significant > normal > unable to assess > missing) will be summarized for each visit, with the.

Descriptive summaries of observed value and change from baseline will be presented for ECG measurements. The number and percentage of patients with QTcB/QTcF>500ms and QTc changes from baseline > 60ms will also be provided. Shift table will summarize change of ECG results from baseline to worst post-baseline value (abnormal with clinically significant > abnormal without clinically significant >normal>missing), and unscheduled visit will be included.

The ECG listing will display observed values and change from baseline. And listings of abnormal results with clinically significant and prolonged QTcF will be provided.

15.6 PHYSICAL EXAMINATION

Physical examination data will be listed. One more listing will be provided including patients with abnormal values only.

16.REFERENCES

1. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika 1988; 75:800-2
2. Guideline on Biostatistics in Drug Clinical Trials (2016 CFDA)
3. Guideline for the Planning and Reporting of Data Management and Statistical Analysis of Drug Clinical Trials (2016 CFDA)
4. ICH E9: Statistical Principles for Clinical Trials 1998

17.APPENDICES

17.1 TIME AND EVENT TABLE

Period	Screening	Baseline	Treatment period				Follow-up/ EW [12]
Visit	1	2	3	4	5	6	7
Study Week	-4 to -1	0	1	4	8	12	14
Study Day	-28to -1	0	7 ± 2 days	28 ± 3 days	56 ± 3 days	84 ± 3 days	98 ± 3 days
Procedures							
Written Informed Consent [1]	x						
Demography/Medical History [2]	x						
Physical Examination	x	x	x	x	x	x	x
Height	x						
Weight	x						x
Chest X-ray[3]	x						
Inclusion/Exclusion Criteria	x						
Randomization criteria		x					
Study Drug Dispensing		x		x	x		
Study Drug Accountability				x	x	x	x [13]
Enrollment Visit in IVRS [4]	x	x	x	x	x	x	x [13]
Assessment of effectiveness							
Investigator's Global Assessment (IGA)	x	x	x	x	x	x	x [13]
Eczema Area and Severity Index (EASI)	x	x	x	x	x	x	x [13]
Scoring Atopic Dermatitis (SCORAD)	x	x	x	x	x	x	x [13]
Pruritus numeric rating scale (NRS)	x	x	x	x	x	x	x [13]
Dermatology quality of life index (DLQI)	x	x	x	x	x	x	x [13]
Photography[5]	x	x	x	x	x	x	x[13]
safety assessment							
12-lead ECG	x	x	x	x	x	x	x

Vital signs (blood pressure, heart rate, temperature and respiratory rate) [6]	x	x	x	x	x	x	x
Adverse Event Assessment [7]		x	x	x	x	x	x
Laboratory Evaluation							
Hematological and biochemical tests	x	x	x	x	x	x	x
Urinalysis	x	x	x	x	x	x	x
Lipid Profile (fasting)		x	x	x	x	x	x
TSH, fT3, fT4	x	x	x	x	x	x	x
QuantiFERON-TB Gold or T-SPOT.TB test	x						
Hepatitis B surface antigen, Hep C Ab, HIV test	x						
Immunoglobulin E (IgE)	x	x	x	x	x	x	x[13]
Urine Pregnancy Test [8]		x	x	x	x	x	x
β-HCG (blood) [9]	x						
[redacted]					[redacted]	[redacted]	
Medication							
Concomitant Medication Assessment	x	x	x	x	x	x	x

1. Written informed consent must be obtained prior to performing any Visit 1 procedures or initiating any alterations in a subject's medications.
2. Atopic Dermatitis Disease History includes collection of details of AD: AD diagnosis, the use of topical treatments, systemic treatments and other treatments for AD.
3. Only to be performed if there is no chest X-ray or CT scan available within 3 months of Visit 1.
4. IVRS is a randomization system used to record all patient visits.
5. Subjects will be photographed at a distance of 1m by a fixed investigator at each visit (the principle of photography is to reflect the characteristics of AD skin lesion to the greatest extent). When photographed, face and private part will be well covered.
6. Vital signs include resting blood pressure, heart rate, body temperature and respiratory rate. It is advised to measure them before any procedures or questionnaires.
7. Adverse events and serious adverse events will be collected from the start of study drug treatment (Visit 2) to the telephone follow-up period. However, any serious adverse events will be recorded from the time of consent.
8. Urine pregnancy test to be done in females of childbearing potential only at scheduled visit
9. β -Human Chorionic Gonadotrophin(β-HCG) to be done in females of childbearing potential at screening visit, and only to be done if urine pregnancy test positive at other study visits.

[redacted]

12. The early withdrawal visit should be done 2 weeks after the last dose/decision of withdrawal.
13. Only to be conducted in the early withdrawal subjects.

17.2 STATISTICAL TABLES, FIGURES AND LISTINGS

The shells for the tables, figures and listings will be provided in a separate document.