Document Type:	Statistical Analysis Plan
Official Title:	A randomized, placebo-controlled, double-blind, parallel-group, multicenter Phase 2a study to investigate efficacy and safety of zabedosertib (BAY 1834845) for the treatment of adult patients with moderate-to-severe atopic dermatitis
NCT Number:	NCT05656911
Document Date:	31-Oct-2023

Title Page

Protocol Title:

A randomized, placebo-controlled, double-blind, parallel-group, multicenter Phase 2a study to investigate efficacy and safety of zabedosertib (BAY 1834845) for the treatment of adult patients with moderate-to-severe atopic dermatitis

Protocol Number:	22158
Compound Number:	Zabedosertib (BAY 1834845)
Short Title:	Proof-of-concept of efficacy and safety of zabedosertib in the treatment of moderate-to-severe atopic dermatitis in adults
Acronym:	Damask
Sponsor Name:	Bayer AG

Legal Registered Address:

Non-US: Bayer AG, 51368 Leverkusen, Germany

US territory: Bayer HealthCare Pharmaceuticals Inc., 100 Bayer Boulevard, P.O. Box 915, Whippany NJ 07981-0915, USA

Regulatory Agency Identifier Number(s):

Registry	ID
EudraCT	2022-000520-38

Date: 31 OCT 2023

Version: 1.0

Statistical Analysis Plan template version: 4.0

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.
The approval of the Statistical Analysis Plan is documented in a separate signature document.

Table of Contents

Table of text tables

Table of text figures

Version History

This Statistical Analysis Plan (SAP) for Study 22158 is based on the protocol Version 2.0 dated 18 JUL 2023.

SAP Version	Date	Change	Rationale
1.0	31 OCT 2023	Not applicable	Original version

List of Abbreviations

AD	atopic dermatitis
AE	adverse event
AESI	adverse event of special interest
ANCOVA	analysis of covariance
BID	twice daily
BMI	body mass index
BSA	body surface area
eCRF	electronic case report form
ClinRO	clinician-reported outcome
EASI	Eczema Area and Severity Index
EOT	end of treatment
FAS	full analysis set
FU	follow up
ICE	intercurrent event
IRS	immunoreactivity score
LOCF	last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
NRS	numerical rating scale
PK	pharmacokinetic
PKS	pharmacokinetic analysis set

PPS	Per protocol analysis set
PRO	patient reported outcome
SAE	serious adverse event
SAF	safety analysis set
SAP	statistical analysis plan
SAS	Statistical Analysis System
SoC	standard of care
TCS	topical corticosteroid
TCI	topical calcineurin inhibitor
TEAE	treatment emergent adverse event
TLF	tables, listings, figures
vIGA-AD	validated Investigator Global Assessment for Atopic Dermatitis
WOCF	Worst observation carried forward

1. Introduction

This statistical analysis plan (SAP) is based on the clinical study protocol Version 2.0 (dated 18-Jul-2023). The SAP describes the administrative interim analysis and final analysis of the study 22158. Table, figure and listing specifications are contained in a separate document.

The following type of data will be included for the study: efficacy, safety, pharmacokinetics and biomarker data.

Changes to the protocol-planned analyses are described in Section 4.8.

For any text copied from the clinical study protocol italic type will be used.

Based on scientific considerations and preclinical observations, zabedosertib may provide a safe and effective treatment option for atopic dermatitis (AD) by its anti-inflammatory properties.

BAY 1834845 is a potent and highly selective interleukin 1 receptor-associated kinase 4 (IRAK4) inhibitor and provides a new approach for the treatment of AD.

IRAK4, is the key isoform of a 4-kinase family regulating the production of proinflammatory cytokines. IRAK4 is the first of the 4 kinases mediating the intracellular signal transduction pathway utilized by the innate immune cell receptors e.g., IL-1R, IL-18 receptor (IL-18R), IL-33 receptor (IL-33R) and Toll-like receptors (TLRs) except TLR3.

The aim of this Phase 2a study is to explore for the first time whether the interleukin 1 receptor-associated kinase 4 (IRAK4) inhibitor zabedosertib can safely provide a therapeutic benefit in patients with moderate-to-severe AD and to achieve a proof of concept for further development steps of the compound (or the compound class).

1.1 Objectives, Endpoints, and Estimands

Objectives	*Endpoints*
Primary objective	*Primary endpoint:*
• *To assess the efficacy of zabedosertib vs. placebo in adult patients with moderate-to-severe atopic dermatitis (AD) with inadequate response to topical corticosteroids or if topical treatments are medically not advisable*	• *Achievement of 75% reduction from baseline in the Eczema Area and Severity Index (EASI 75 response) at Week 12 (Day 84)* In detail, *primary endpoint (*see below*) is the composite variable defined as follows:* • *an EASI 75 response at Week 12 (Day 84),* • *no stop of study intervention for reasons related to lack of efficacy,* • *no rescue medication use during the 4 weeks before Day 84 and* • *no use of systemic AD treatment.*
	*Secondary endpoints (*for the primary objective*):* • *Percent change from baseline in EASI at Week 12 (Day 84)* • *Achievement of EASI 50 response at Week 12 (Day 84)* • *Achievement of EASI 90 response at Week 12 (Day 84)* • *Achievement of a vIGA-AD response (score 0 or 1 and ⩾ 2 points improvement) at Week 12 (Day 84)* • *Absolute change from baseline in body surface area (BSA) affected by AD at Week 12 (Day 84)* • *Absolute values and percent change of the weekly average of the Peak Pruritus 0-10 NRS score from baseline at Week 12 (Day 84)* • *Achievement of a ⩾ 4 point-improvement (reduction) in the weekly average of the Peak Pruritus 0-10 NRS score from baseline at Week 12 (Day 84) for participants with Peak Pruritus 0-10 NRS score ⩾ 4 at baseline*
Secondary objectives	
• *To assess the safety and tolerability of zabedosertib vs. placebo*	*Secondary endpoints:* • *Frequency and severity of TEAEs* *Other safety endpoints:* • *Frequency and severity of TEAEs of special interest* • *Changes in vital signs from baseline* • *Changes in clinical laboratory test results from baseline (e.g., chemistry lipid panel, hematology, urinalyses)* • *Changes in ECG parameters from baseline and frequency of ECG findings*
Other pre-specified / Exploratory objectives	*Other Pre-Specified Endpoints*
• *To further assess the efficacy of zabedosertib vs. placebo*	• *Achievement of EASI 75 response up to Week 8 (Day 56)* • *Achievement of EASI 100 response at Week 12 (Day 84)* • *Achievement of a ⩾ 3 point- improvement (reduction) in the weekly average of the Peak Pruritus 0-10 NRS from baseline at Week 12 (Day 84)* • *Absolute values and percent change from baseline in weekly averages of Peak Pruritus 0-10 NRS scores over time*
• *Exploration of the pharmacokinetics of zabedosertib*	• *Plasma concentrations of zabedosertib and metabolite BAY 28822815 (M-5) and population pharmacokinetic (popPK) evaluation derived PK parameters (will be reported separately, if applicable)*
• *Pharmacodynamics disease-monitoring biomarkers by:*	

Objectives	*Endpoints*
• *Assessment of treatment response by analysis of blood biomarkers (BM)* • *Evaluation of skin biopsies (subset of participants)*	• *Plasma levels of inflammatory markers (optional analysis)* • *Measurement of molecular changes in AD transcriptome reg. inflammatory cytokines, chemokines (e.g., IL-1β, IL-4, IL-33, TSLP) and immunohistochemistry (e.g., epidermal barrier markers; optional analysis).*

The estimand for the primary endpoint (analysis on PPS) is shown below.

Primary Objective: *See* above
Estimand label: *primary*
Endpoint: *Achievement of 75% reduction from baseline in the Eczema Area and Severity Index (EASI) at Week 12 (Day 84) – EASI 75 response*
Population: *Adult patients with moderate-to-severe atopic dermatitis (for at least one year and inadequate response to topical treatment or if topical treatments are medically not advisable)*
Treatment condition: *Zabedosertib vs placebo*
Population-level summary: *The estimated difference in the proportion of responders between interventions.*
Intercurrent event(s) • *Use of topical rescue medication:* *At Day 56 (Visit 6) or later* (until Day 84 (Visit 7) or end of treatment if before Day 84 (Visit 7)): *analyze study participant as non-responder (composite strategy)* *Before Day 56 (Visit 6):* *disregard use of topical rescue medication (treatment policy strategy)* • *Use of systemic standard of care:* *analyze study participant as non-responder (composite strategy)* (SoC medication use after early discontinuation due to other reasons will not trigger the classification as non-responder for the following time points). • *Discontinuation of treatment due to lack of efficacy:* *analyze study participant as non-responder (composite strategy)* • *Discontinuation of treatment due to reasons not related to lack of efficacy (including Covid-19 related reasons):* *At Day 56 (Visit 6) or later:* *imputation of response information (hypothetical strategy)* *Before Day 56 (Visit 6):* *exclusion from analysis as adequate imputation of response information is considered not possible* • *Non-compliance with study intervention (general), defined as actual drug intake lower than 80% of the planned drug intake. Actual drug intake will be calculated from the date of first drug intake until end of treatment or start of rescue medication (at or after Day 56 (Visit 6)), whatever comes first:* *exclusion from analysis as adequate imputation of response information is considered not possible* • *Non-compliance with study intervention (last 4 weeks) defined as actual drug intake lower than 80% of the planned drug intake during the last 4 weeks. Reference point for determination of the time period will be either the end of treatment or the start of rescue medication (at or after Day 56*

(Visit 6)), whatever comes first:
exclusion from analysis as adequate imputation of response information is considered not possible

- *Non-compliance with emollients / moisturizers:*
  regardless of non-compliance with emollients / moisturizers (treatment policy)
- *Missing baseline value EASI assessment:*
  exclusion from analysis as adequate imputation of response information is considered not possible
- *Missing Day 84 (Visit 7) EASI assessment:*
  imputation of response information using data from Day 56 (Visit 6) or later (hypothetical strategy), if none of the conditions applies which are described above.
  (If an imputation based on Day 56 (Visit 6) is not possible due to scarcity of data, imputation might be based on data from earlier time points or baseline only.)
- *Missing Day 56 (Visit 6) and later EASI assessment:*
  exclusion from analysis unless participant is a non-responder due to use of rescue medication, SOC or treatment discontinuation due to lack of efficacy

In the following a further exploratory estimand is listed. Changes to the primary estimand are printed in bold. This estimand will be used for analyses of the primary endpoint on the FAS.

Primary Objective: *See* above
Estimand label: supplementary - FAS
Endpoint: *Achievement of 75% reduction from baseline in the Eczema Area and Severity Index (EASI) at Week 12 (Day 84) – EASI 75 response*
Population: *Adult patients with moderate-to-severe atopic dermatitis (for at least one year and inadequate response to topical treatment or if topical treatments are medically not advisable)*
Treatment condition: *Zabedosertib vs placebo*
Population-level summary: *The estimated difference in the proportion of responders between interventions.*
Intercurrent event(s) • *Use of topical rescue medication:* *At Day 56 (Visit 6) or later* (until Day 84 (Visit 7) or end of treatment if before Day 84 (Visit 7))*:* *analyze study participant as non-responder (composite strategy)* *Before Day 56 (Visit 6):* *disregard use of topical rescue medication (treatment policy strategy)* • *Use of systemic standard of care:* *analyze study participant as non-responder (composite strategy)* (SoC medication use after early discontinuation due to other reasons will not trigger the classification as non-responder for the following time points.) • *Discontinuation of treatment due to lack of efficacy:* *analyze study participant as non-responder (composite strategy)* • *Discontinuation of treatment due to reasons not related to lack of efficacy (including Covid-19 related reasons):* **imputation of response information (hypothetical strategy)**

- *Non-compliance with emollients / moisturizers:*
  regardless of non-compliance with emollients / moisturizers (treatment policy)
- *Missing baseline value EASI assessment:*
  exclusion from analysis as adequate imputation of response information is considered not possible
- *Missing Day 84 (Visit 7) EASI assessment:*
  imputation of response information (hypothetical strategy) if none of the conditions applies which are described above
- *Missing Day 56 (Visit 6) and later EASI assessment:*
  imputation of response information (hypothetical strategy) if none of the conditions applies which are described above

Primary Objective: See above

Estimand label: exploratory – all post-dose time points (excluding FU visit)

Endpoint:
Achievement of 75% reduction from baseline in the Eczema Area and Severity Index (EASI) over time – EASI 75 response

Population:
Adult patients with moderate-to-severe atopic dermatitis (for at least one year and inadequate response to topical treatment or if topical treatments are medically not advisable)

Treatment condition:
Zabedosertib vs placebo

Population-level summary:
The estimated difference in the proportion of responders between interventions.

Intercurrent event(s)

- Use of topical rescue medication (until Day 84 (Visit 7) or end of treatment if before Day 84 (Visit 7)):
  analyze study participant as non-responder (composite strategy)
- Use of systemic standard of care:
  analyze study participant as non-responder (composite strategy)
  (SoC medication use after early discontinuation due to other reasons will not trigger the classification as non-responder for the following time points.)
- Discontinuation of treatment due to lack of efficacy:
  analyze study participant as non-responder (composite strategy)
- Discontinuation of treatment due to reasons not related to lack of efficacy (including Covid-19 related reasons):

  PPS:
  At Day 56 (Visit 6) or later:
  imputation of response information (hypothetical strategy)

  Before Day 56 (Visit 6):
  exclusion from analysis as adequate imputation of response information is considered not possible
  imputation of response information (hypothetical strategy)

  FAS:
  imputation of response information (hypothetical strategy)

- Non-compliance with emollients / moisturizers:
  regardless of non-compliance with emollients / moisturizers (treatment policy)
- Missing baseline value EASI assessment:
  exclusion from analysis as adequate imputation of response information is considered not possible
- Missing EASI assessment:
  imputation of response information (hypothetical strategy) if none of the conditions applies which are described above

1.2 Study Design

The study population is characterized by the following main inclusion criteria:

- 18 to 65 years of age inclusive
- Diagnosis of AD for ≥ 1 year
- Moderate-to-severe AD at randomization visit as defined by
  - EASI score ≥ 16,
  - BSA affected by AD ≥ 10%,
  - vIGA-AD score ≥ 3, and
  - Peak Pruritus 0-10 NRS ≥ 4 (average score of the daily scores of the 7 days before randomization, with ≥ 4 scores required).
- Documented history (within 6 months prior to the first screening visit) of inadequate response to treatment with topical corticosteroids (TCS), or if TCS are medically not advisable (e.g., due to important side effects or safety risks)
- Stable amount of emollient applied to skin over the whole body twice daily for at least the 7 consecutive days before the randomization visit
- Body mass index (BMI) within the range of 18.5 to 35.0 kg/m2 (inclusive)

The design of the study is as follows:

Study phase:	Phase 2a (Proof of concept)
Study population:	Adult patients with moderate-to-severe atopic dermatitis
Design:	Parallel group design
Blinding:	Double-blind
Placebo controlled:	Yes
Randomized:	Yes
Strata:	No
Countries:	7 countries, multicenter
Treatment:	Zabedosertib 120 mg
	Placebo
Frequency of treatment:	Twice daily

Route of treatment: Oral intake

Duration of treatment: 12 weeks (84 days)

A total of 72 participants are planned to be randomly allocated to 1 of the 2 treatment groups with a 2:1 allocation ratio to achieve a total of 57 participants valid for efficacy analysis:

zabedosertib at a dose of 120 mg BID:	*48 randomized / 38 valid for the efficacy analysis*
Matching placebo BID:	*24 randomized / 19 valid for the efficacy analysis.*

The study consists of

- *a 28-day screening period (Visits 1 and 2),*
- *a 12-week intervention period consisting of 5 visits (Visits 3, 4, 5, 6, and 7),*
- *and a 4-week follow-up (FU) period with one visit (Visit 8).*

Thus, the total study duration per participant will be 17 to 20 weeks (approximately 140 days), with 8 study visits at intervals of 8 to 28 ± 3 days.

The primary analysis will be performed after database release.

2. Statistical Hypothesis

The primary efficacy endpoint is achievement of 75% reduction from baseline in the EASI (EASI 75 response) at Week 12 (Day 84, EOT).

Let π_{BAY} and π_{Plac} be the responder rate with respect to EASI 75 after zabedosertib and placebo, respectively.

The statistical hypothesis is defined as follows:

$\pi_{BAY} > \pi_{Plac}$ *('clinical activity' criterion),*

i.e. that the EASI 75 responder rate after treatment with zabedosertib is higher than after placebo.

The study will be evaluated using Bayesian statistics, which allows to determine the posterior probability in favor of a hypothesis. Therefore, it is not needed to specify a pair of hypotheses. The statistical hypothesis will be considered accepted if the posterior probability for the hypothesis of interest is higher than 90%.

2.1 Multiplicity Adjustment

No adjustments for multiplicity are planned since this is an exploratory study.

3. Analysis Sets

For the purpose of analysis, the following analysis sets are defined:

Analysis set	***Description***
Enrolled	*All participants who signed the ICF.*
All Randomized *(RAND)*	*All participants randomly assigned to study intervention* *Participants will be analyzed according to the intervention they were randomized to.*
Safety analysis set *(SAF)*	*All participants who took at least 1 dose of study intervention.* *Participants will be analyzed according to the intervention they actually received.*
Per protocol set *(PPS)*	*All randomized participants who meet the following criteria:* – *have an EASI score at baseline* – *have an EASI score at Day 56 (Visit 6) or later or took rescue medication between Day 56 (Visit 6) and Day 84 (Visit 7) or took SoC* (see Section 8.3 for definition of SoC) *or discontinued study intervention due to lack of efficacy (at any time)* (SoC medication use after early discontinuation due to other reasons will not qualify for the PPS) – *showed compliance of at least 80% with study intervention between start of treatment and end of treatment or start of rescue medication at or after Day 56 (Visit 6), whatever comes first* – *showed compliance of at least 80% with study intervention during the last 4 weeks before end of treatment or start of rescue medication at or after Day 56 (Visit 6), whatever comes first.* – *are without validity findings with respect to the efficacy related entry criteria* *Participants will be analyzed according to the treatment they actually received.*
Pharmacokinetic set (PKS)	*All participants who* – *were treated with zabedosertib,* – *have at least 1 valid zabedosertib plasma concentration sample,* – *are without protocol deviations, which would interfere with the evaluation of the PK data.*
Full Analysis Set (FAS)	All participants who – have an EASI score at baseline, – have a valid EASI score at Day 14 (Visit 4) or later or took rescue medication between Day 56 (Visit 6) and Day 84 (Visit 7) or took SoC (see Section 8.3 for definition of SoC) or discontinued study intervention due to lack of efficacy (at any time), (SoC medication use after early discontinuation due to other reasons will not qualify for the FAS) – took at least 1 dose of study intervention, – are without validity findings with respect to the main efficacy related entry criteria. Participants will be analyzed according to the treatment they actually received.
Biomarker analysis set (BAS)	All participants who – showed compliance of at least 80% with study intervention during the last 4 weeks before end of treatment – have at least 1 valid biomarker sample at baseline and post-dose (blood sample or biopsy), – are without validity findings, which would interfere with the evaluation of the biomarker data, – are without validity findings with respect to the main efficacy related entry criteria Participants will be analyzed according to the treatment they actually received.

Final decisions regarding the assignment of participants to analysis sets will be made during the blind review of study data and documented in the final list of important deviations, validity findings and assignment to analysis set(s). Any case-by-case decision concerning

validity will be taken during the blind review by persons without access to any individual post-dose efficacy data to avoid any potential bias.

Please also refer to Section 8.3 for the assessment of the intercurrent events “use of rescue medication”, “use of systemic standard of care (SoC)”, and “discontinuation of study intervention due to lack of efficacy” and Section 4.5.1 for the calculation of compliance used in the definition of the analysis sets FAS and PPS.

4. Statistical Analyses

4.1 General Considerations

4.1.1 General principles and data rules

The statistical evaluation will be performed by using the software SAS (release 9.4 or higher; SAS Institute Inc., Cary, NC, USA).

Continuous variables assumed to be normally distributed will be summarized using number of participants, mean, standard deviation, minimum, median, and maximum. Continuous variables assumed to be log-normally distributed will be summarized using number of participants, geometric mean, geometric standard deviation, geometric coefficient of variation, arithmetic mean, arithmetic standard deviation, arithmetic coefficient of variation, minimum, median, and maximum.

For categorical data frequency tables will be provided.

In general, summary statistics and frequency tables will be presented by treatment group, visit and time point (if applicable). Plots over time (if applicable) will also be presented by treatment group.

The baseline assessment will be the assessment at Day 1 (Visit 3). For all variables with exception of the efficacy variables (ClinRO and PRO data) the following applies: If the assessment on Day 1 (Visit 3) is missing, then baseline will be the last available assessment before first study drug intake (i. e. usually the screening assessment).

Summary statistics of continuous variables will be provided including absolute changes from baseline or percent changes from baseline. Absolute changes from baseline will be calculated as (post-dose value - baseline value) unless noted otherwise. Percent changes from baseline will be calculated as ((post-dose value - baseline value)/baseline value)*100 unless noted otherwise.

Efficacy variables at FU visits will only be evaluated by summary statistics, but they will not be included for statistical analyses such as Bayesian analyses, ANCOVA models and mixed models for repeated measures.

In case of replicate measurements for one visit mean values of the unrounded original values will be calculated by participant and visit. Summary statistics of these mean values will be provided.

Box plots will be displayed with whiskers of maximum 1.5 times the interquartile range (i. e. upper quartile minus lower quartile) above the upper quartile, and a distance of maximum 1.5 times the interquartile range below the lower quartile. All data points above or below the whiskers will be displayed as single data points.

For analytes which are below the lower limit of quantification (LLOQ), one half of the LLOQ will be used as an estimate for the respective value. However, means, standard deviations and coefficients of variation will only be calculated if at least 2/3 of the values are above (or equal to) the LLOQ. In tables showing mean values, where values below LLOQ are included in the calculation of mean values, these means will be marked. Corresponding tables and figures will get a footnote indicating that "Values below the lower limit of quantification were replaced by ½ LLOQ if at least 2/3 of the values were >= LLOQ." Tables displaying minimum values will present "<LLOQ" as minimum. Similar rules will apply to data points above ULOQ. ULOQ will be used as an estimate for the respective values. Means, standard deviations and coefficients of variation for changes to baseline will only be calculated if there are at least 2/3 of the changes where the baseline and post-dose value are >= LLOQ and <=ULOQ.

The relative onset of adverse events with respect to treatment start will be expressed in days as follows (similar for relative end of adverse events):

- date of adverse event start minus date of treatment start for adverse events starting before treatment start
- date of adverse event start minus date of treatment start + 1 for adverse events starting on day of first treatment or later.

The duration of an AE will be calculated in days, i. e., the date of adverse event end minus the date of adverse event start +1.

The relative onset of prior/concomitant/posterior medication with respect to treatment start will be expressed in days as follows (similar for relative end of prior/concomitant/posterior medication):

- date of medication start minus date of treatment start for medications starting before treatment start
- date of medication start minus date of treatment start + 1 for medications starting on day of treatment start or later.

The duration of atopic dermatitis (in years) at baseline will be calculated as follows:

- (date of Day 1 (Visit 3) minus date of first diagnosis plus 1) / 365.25.

The age at onset (in years) will be calculated as follows:

- year of AD onset – year of birth.

Data from unscheduled measurements will be listed, but not included in summary statistics or frequency tables by visit unless specified otherwise (refer to Section 8.4 for handling of unscheduled measurements for efficacy variables).

Listings will be provided sorted by treatment group and participant identifier ('participant' and 'subject' will be used interchangeable in the tables, listings and figures and the CSR).

The following labels will be used for TLF outputs (unless indicated otherwise):

- Treatment groups:

  Zabedosertib

  Placebo,

- Visits:

  Day -28 (Visit 1)

  Day -8 (Visit 2)

  Day 1 (Visit 3)

  Day 14 (Visit 4)

  Day 28 (Visit 5)

  Day 56 (Visit 6)

  Day 84 (Visit 7)

  FU visit.

All efficacy analyses will be performed on the PPS. In addition, the efficacy analyses will be performed on the FAS - unless otherwise indicated. Participants will be analyzed according to the treatment they actually received.

All safety analyses will be performed on the SAF, participants will be analyzed according to the treatment they actually received.

4.1.2 Handling of missing data

Participants who discontinue the study prematurely (i. e. drop out) will not be replaced.

All data will be presented in the participant data listing as they are recorded on the Case Report Form (CRF), i. e., partially missing data will appear as such.

For participants who permanently discontinue the study intervention, an EOT visit (…) *should be performed as soon as the study intervention has been discontinued. If the decision to stop the study intervention early is taken during a site visit, this visit becomes the EOT visit. The FU visit should be performed as planned after the last dose.*

In general, data collected at early discontinuation visits will be handled as follows:
If a valid value is available from the EOT visit and if the assessment time is in accordance with one of the upcoming visit windows (see Section 8.4.1 for visit windows) and is not assessed later than three days after last study drug intake, this value will be assigned to the respective visit for analysis. Values assessed at FU visits after EOT visits due to early discontinuation will be summarized together with FU visits after regular EOT visits at Day 84 (Visit 7) unless specified otherwise.

The data handling of efficacy data which might deviate from the procedure described above depending on intercurrent events is described in Section 8.4.

If parts of the start or stop date of an AE are missing (day, month or year) the relative onset of the AE with respect to treatment start and/or the duration of the adverse event will not be calculated.

Adverse events will be considered as treatment-emergent if they occur after the first study intervention intake and until 7 (calendar) days after last study drug intake.

To determine if an AE is treatment-emergent in case of missing or incomplete dates the following rules will be applied considering a "worst case" scenario, i. e. if it is possible that the AE is treatment emergent it will be considered as treatment-emergent:

- If the date of AE start is completely missing, the AE will be considered as treatment-emergent unless the stop date (or other information) suggests otherwise
- If the day of AE start is missing, but the month and year is known, the AE will be considered as treatment-emergent if the month of AE start is between the month of treatment start and the month of treatment end date plus 7 days (both included), unless the stop date of the AE (or other information) suggests otherwise
- If the day and month of AE start are missing, but the year is known, the AE will be considered as treatment-emergent if the year of AE start is between the year of treatment start and the year of treatment end date plus 7 days (both included), unless the stop date (or other information) suggests otherwise.

To determine if a medication is prior, concomitant or posterior with respect to study intervention intake in case of missing or incomplete dates a "worst case" scenario will be applied, i. e. if it is possible that a medication is prior it will be considered prior, if it is possible that a medication is concomitant it will be considered concomitant, and if its possible that a medication is posterior it will be considered posterior.

To determine if a procedure is prior, concomitant or posterior with respect to study intervention intake in case of missing or incomplete dates a "worst case" scenario will be applied, i. e. if it is possible that a procedure is prior it will be considered prior, if it is possible that a procedure is concomitant it will be considered concomitant, and if its possible that a procedure is posterior it will be considered posterior (for definition of prior, concomitant and posterior medications and procedures, refer to Section 6.3).

Duration of atopic dermatitis will be calculated as follows if date parts of first diagnosis are missing: if the day and month of first diagnosis are missing the day will be imputed as 1st of July, if only the day of first diagnosis is missing it will be imputed as the 15th of the month.

4.2 Primary Endpoint Analysis

4.2.1 Definition of Endpoint(s)

The primary efficacy endpoint is achievement of 75% reduction from baseline in the Eczema Area and Severity Index (EASI 75 response) at Week 12 (Day 84).

The EASI score ranges from 0 to 72. The lower the EASI score the better the state of the participant with respect to severity and extent of AD symptoms.

The EASI is a ClinRO assessing the extent of AD at four body regions (head and neck, trunk and upper and lower extremities) by measuring the average severity of four clinical signs at each body region: (1) erythema, (2) induration/papulation, (3) excoriation, and (4) lichenification, each on a scale of 0 to 3 (Hanifin et al, 2001).

For scoring, the sum of these severity scores for each region is multiplied by a region score, a value of between 0 and 6 representing the percentage of skin affected by eczema, and finally weighted by a body region dependent multiplier (0.1 for head/neck; 0.3 for trunk, 0.2 for

upper extremities and 0.4 for lower extremities). This provides an EASI score per body region, with the final EASI score being the sum of the 4 region scores. The minimum EASI score is 0 and the maximum EASI score is 72, with a higher score indicating worse severity of AD. Further information and a scoring table can be found in Section 8.1.1.

4.2.2 Main Analytical Approach

The analysis of the primary efficacy endpoint on the PPS will be performed as follows (according to the primary estimand, refer to Section 1.1)*:*

The impact of a potential use of rescue medication (TCSs or TCIs) or SoC for treatment of AD on the efficacy effect will be accounted for as follows:

If rescue medication is used between Day 56 (Visit 6) and Day 84 (Visit 7) (or end of treatment if before Day 84 (Visit 7)) *or if SoC is used (at any time), the participant will be specified as a non-responder with respect to the EASI 75 response from the time the rescue medication or SoC medication is used.* SoC medication use after early discontinuation due to other reasons not related to AD will not trigger the classification as non-responder for the following time points.

If a participant discontinues study intervention due to lack of efficacy at any time (e.g., due to an adverse event related to AD such as disease flare), the participant will also be specified as a non-responder with respect to EASI 75 response from the time of the discontinuation.

Otherwise, missing data - e.g. due to study intervention discontinuation due to other reasons such as reasons related to Covid-19 pandemic (individual reasons or prevention/logistic measures), withdrawal of consent not due to lack of efficacy or adverse events not related to AD - will be imputed using a multiple imputation approach based on logistic regression model.

Bayesian inference will be used to quantify the EASI 75 responder rates on Day 84 in the zabedosertib and placebo treatment group. For each treatment group a non-informative Beta(1,1) prior will be used. The Beta(1,1) prior produces a posterior distribution in the same distribution family, namely Beta(k+1,n-k+1) with k being the number of EASI 75 responders in the respective treatment group and n being the number of all participants in the respective treatment group. Simulations based on the posterior distributions of the EASI 75 response rate in both treatment groups will then be used to calculate the posterior probability $P(\pi_{BAY} > \pi_{Plac} \mid data)$*. In addition, median and* 80% *and 95% credible intervals will be calculated based on the posterior distributions for the response rates in both treatment groups as well as for the difference between both treatment groups.*

Code similar to the following will be used for the Bayesian analysis:

```
data test;
      call streaminit(1234);
      do i=1 to &simno.;
            *simulation BAY;
            p_bay=rand("beta", &k_bay.+1,&n_bay.-&k_bay.+1);
            *simulation Placebo;
            p_pla=rand("beta", &k_pla.+1,&n_pla.-&k_pla.+1);
            *counts for advantage for zabedosertib;
            diff = p_bay - p_pla;
```

```
            if diff > 0 then count=1;
            else count=0;
            output;
      end;
run;
```

A seed of 1234 will be used and at least 100000 iterations will be performed.

The posterior probability for $P(\pi_{BAY} > \pi_{Plac} \mid data)$ will then be calculated as mean of `count`. The posterior distribution of the difference between the frequency of responders in the two treatment groups (`diff`) will be summarized by median, 2.5th and 97.5th percentile (95% credible interval) as well as 10th and 90th percentile (80% credible interval). The posterior distribution for the frequency of responders in the zabedosertib (`p_bay`) and placebo (`p_pla`) group will also be summarized by median, 2.5th and 97.5th percentile (95% credible interval) as well as 10th and 90th percentile (80% credible interval).

Histograms will be presented for the posterior distribution of achievement of EASI 75 response at Day 84 (Visit 7) for zabedosertib, placebo and the difference (zabedosertib – placebo) in Section 10 of the CSR.

Imputation of missing EASI 75 response information for Day 84 (Visit 7)

For a detailed technical description and data handling steps for the imputation of missing EASI 75 response information please also refer to Section 8.4.1 and 8.4.2. In the following the imputation procedure will be described in general:

In a first step imputation of missing EASI 75 response information will be done for all post-dose visits according to the primary estimand as follows:

- *Use of topical rescue medication:*
  At Day 56 (Visit 6) or later (until Day 84 (Visit 7) or end of treatment if before Day 84 (Visit 7))*:*
  analyze study participant as non-responder (composite strategy)

  Before Day 56 (Visit 6):
  disregard use of topical rescue medication (treatment policy strategy)
- *Use of systemic standard of care:*
  analyze study participant as non-responder (composite strategy)
  (SoC medication use after early discontinuation due to other reasons not related to AD will not trigger the classification as non-responder for the following time points.)
- *Discontinuation of treatment due to lack of efficacy:*
  analyze study participant as non-responder (composite strategy)

In the next step, imputation of still missing EASI 75 response information at Day 84 (Visit 7) will be done (in case of discontinuation of treatment due to reasons not related to lack of efficacy at Day 56 (Visit 6) or later or a missing Day 84 (Visit 7) EASI score) according to the primary estimand:

- *Discontinuation of treatment due to reasons not related to lack of efficacy (including Covid-19 related reasons):*
  At Day 56 (Visit 6) or later:
  imputation of response information (hypothetical strategy)
- *Missing Day 84 (Visit 7) EASI assessment:*
  imputation of response information using data from Day 56 (Visit 6) or later (hypothetical strategy), if

none of the conditions applies which are described above.
(If an imputation based on Day 56 (Visit 6) is not possible due to scarcity of data, imputation might be based on data from earlier time points or baseline only.)

In detail, the imputation will be done for the following participants in the PPS:

- with EASI score at baseline and Day 56 (Visit 6),
- without EASI 75 response information at Day 84 (Visit 7),
- without systemic SoC use before Day 56 (Visit 6) EASI score,
- without rescue medication use at Day 56 (Visit 6) EASI score.

The imputation for the primary analysis will be performed by treatment group using the monotone logistic regression method from the SAS multiple imputation procedure (PROC MI). A logistic regression model will be built based on the following participants in the PPS:

- with EASI score at baseline and Day 56 (Visit 6),
- with EASI 75 response information at Day 84 (Visit 7),
- without systemic SoC use before Day 56 (Visit 6) EASI score,
- without rescue medication use at Day 56 (Visit 6) EASI score,
- without rescue medication use (irrespectively of the indication) within the 7 days before the Day 56 (Visit 6) EASI score (to avoid influence of EASI scores directly after rescue medication on the imputation of the primary endpoint).

Participants with intercurrent event use of systemic SoC or rescue medication or discontinuation due to lack of efficacy after the Day 56 (Visit 6) EASI score and before Day 84 (Visit 7) will have EASI 75 response information (non-response) at Day 84 (Visit 7) and can be included in the logistic regression model if they fulfill the criteria specified above.

The EASI 75 response information at Day 84 (Visit 7) (coding: responder=1, non-responder=0) will be the dependent variable in the logistic regression model and the following variables will be included as independent variables:

- EASI score at baseline,
- EASI score at Day 56 (Visit 6).

If the multiple imputation using logistic regression as described above is based on less than 8 participants per treatment group the imputation with logistic regression as described above will be performed with the EASI score at baseline and at the previous visit as independent variable. If there will not be sufficient participants to include one of the post-dose visits (Day 28 (Visit 5), or – if not possible – Day 14 (Visit 4)) the logistic regression model will only use the EASI score at baseline as independent variable. Then the model will be based on all participants with EASI score at baseline and EASI 75 response at Day 84 (Visit 7) which might be imputed as non-response after SoC use, discontinuation of study intervention due to lack of efficacy or use of rescue medication after Day 56 (Visit 6). The variables used in the regression model will be indicated below the output table.

If there is a participant with rescue medication use within the 7 days before the Day 56 (Visit 6) EASI score and missing EASI 75 response at Day 84 (Visit 7) which needs to be imputed

the same procedure as described above (in case of less than 8 particpants per treatment group) will be applied for that participant.

Imputation according to the procedure defined above will be performed 50 times. All participants with values imputed using logistic regression will be combined with the remaining participants (which will have the same values for each imputed dataset).

Bayesian analysis will then be performed based on all participants in the PPS as described above for each imputed dataset to generate posterior samples for the frequency of responders in the zabedosertib (p_bay) and placebo (p_pla) group and the difference (diff) between both.

As suggested by Zhou and Reiter (2010) the Bayesian posterior probability that the EASI 75 responder frequency after treatment with zabedosertib is higher than after placebo will then be derived based on the combined simulated posterior distribution (i. e. combining the posterior samples from all imputed datasets) of the difference in EASI 75 responder frequencies (zabedosertib – placebo). Medians, 80% and 95% credible intervals and histograms will be calculated based on the combined simulated posterior distribution as well. The same will be done for the EASI 75 responder frequencies in the zabedosertib and placebo group.

4.2.3 Sensitivity Analyses

A sensitivity analysis will be performed with respect to imputation related to the following intercurrent event to assess the robustness of the analysis with regard to missing data:

- *Discontinuation of treatment due to reasons not related to lack of efficacy (including Covid-19 related reasons):*
  At Day 56 (Visit 6) or later:
  imputation of response information (hypothetical strategy)
- *Missing Day 84 (Visit 7) EASI assessment imputation of response information using data from Day 56 (Visit 6) or later (hypothetical strategy), if none of the conditions applies which are described above*

As a “worst case” scenario the primary analysis (according to primary estimand) will be repeated by imputing the missing EASI 75 response information at Day 84 (Visit 7) as non-response (instead of using the multiple imputation with the logistic regression model).

Furthermore, the primary analysis (according to primary estimand) will be repeated by imputing the missing EASI 75 response information at Day 84 (Visit 7) using last observation carried forward (instead of using the multiple imputation with the logistic regression model) (refer to Section 8.4.2 for details). EASI scores will only be carried forward if no rescue medication was used within the last 7 days before the EASI score to avoid influence of EASI scores directly after rescue medication on the imputation of the primary endpoint.

4.2.4 Supplementary Analyses

Handling of use of rescue medication

The following supplementary analyses will be performed on the PPS to further assess the influence of rescue medication use on the EASI 75 response.

The analysis will be performed with a different handling of the following intercurrent event:

- *Use of topical rescue medication:*
  At Day 56 (Visit 6) or later (until Day 84 (Visit 7) or end of treatment if before Day 84 (Visit 7))*:*
  analyze study participant as non-responder (composite strategy)

  Before Day 56 (Visit 6):
  disregard use of topical rescue medication (treatment policy strategy)

Use of rescue medication will be handled for supplementary analyses as outlined in the following:

The study participant is considered a non-responder in case of rescue medication use between the EOT visit and

- *Baseline,*
- *2 weeks after baseline,*
- *4 weeks after baseline.*

One additional analysis will be performed where rescue medication does not trigger the classification as non-responder at all to assess the effect of combined treatment of zabedosertib with topical AD medication.

If rescue medication does not trigger the classification as non-responder, the EASI 75 response will be based on the EASI scores collected after use of rescue medication.

Participants who take SoC medication or who discontinue study intervention due to lack of efficacy will still be counted as a non-responder for the time points after SoC use or discontinuation of study intervention due to lack of efficacy as described above.

The EASI score at time points after discontinuation of treatment due to other reasons not related to lack of efficacy or otherwise missing values will be imputed using “worst case” and last observation carried forward as described in Section 4.2.3. No multiple imputation approach using logistic regression will be performed for the supplementary analysis due to potential scarceness of participants continuing until Day 56 (Visit 6) without any rescue medication.

All other intercurrent events will be handled as described for the primary estimand.

Analysis on FAS

The EASI 75 response will be analyzed on the FAS according to the supplementary estimand (refer to Section 1.1). Imputation techniques in case of missing EASI 75 response information at Day 84 (Visit 7) will be as for the primary analysis.

However, the missing EASI 75 response information at Day 84 (Visit 7) due to other reasons not related to lack of efficacy will only be imputed as non-response as well as using last

observation carried forward as described in Section 4.2.3. No multiple imputation approach using logistic regression will be performed to impute EASI score at Day 84 (Visit 7) for the FAS due to potential scarceness of participants without use of rescue medication in the 7 days before EASI scores at early time points before Day 56 (Visit 6).

The supplementary analyses with respect to use of rescue medication as outlined above will not be performed on the FAS with exception of the following:

The analyses on the FAS as described above will be repeated with rescue medication use triggering the classification of non-response for all subsequent visits irrespectively of the time of rescue medication use (i. e. starting from Day 1 (Visit 1) until Day 84 (Visit 7) or end of treatment.

Table 4-1 summarizes the analyses for the primary endpoint EASI 75 response at Day 84 (Visit 7):

Table 4-1: Primary, sensitivity and supplementary analyses for the primary endpoint EASI 75 response at Day 84 (Visit 7)

Analysis set	Analysis type	After "Use of rescue medication"	After "Use of SoC" / "discontinuation due to lack of efficacy"	After "Discontinuation due to other reasons" or missing values	No. of binary analysis
PPS	Primary	Non-response at/after Day 56 (Visit 6)	Non-response	Multiple imputation using logistic regression	B_1.1 (Primary)
	Sensitivity	Non-response at/after Day 56 (Visit 6)	Non-response	Non-response	B_1.2 (Sensitivity 1.1)
				Last observation carried forward	B_1.3 (Sensitivity 1.2)
	Supplementary	Non-response after Day 1 (Visit 3)	Non-response	Non-response	B_2.1 (Supplementary 1.1)

				Last observation carried forward	B_2.2 (Supple-mentary1 .2)
		Non-response at/after Day 14 (Visit 4)	Non-response	Non-response	B_3.1 (Supple-mentary 2.1)
				Last observation carried forward	B_3.2 (Supple-mentary 2.2)
		Non-response at/after Day 28 (Visit 5)	Non-response	Non-response	B_4.1 (Supple-mentary 3.1)
				Last observation carried forward	B_4.2 (Supple-mentary 3.2)
		Not considered as non-response at any time	Non-response	Non-response	B_5.1 (Supple-mentary 4.1)
				Last observation carried forward	B_5.2 (Supple-mentary 4.2)
FAS	Supplementary	Non-response at/after Day 56 (Visit 6)	Non-response	Non-response	B_6.1 (Supple-mentary 5.1)

				Last observation carried forward	B_6.2 (Supplementary 5.2)
		Non-response after Day 1 (Visit 3)	Non-response	Non-response	B_7.1 (Supplementary 6.1)
				Last observation carried forward	B_7.2 (Supplementary 6.2)

4.3 Secondary Endpoint(s) Analysis

4.3.1 Key Secondary Endpoint(s)

Not applicable.

4.3.2 Supportive Secondary Endpoint(s)

Calculation of the weekly average of Peak Pruritus 0-10 NRS

The weekly average of the Peak Pruritus 0-10 NRS will be evaluated in a visit-based manner, in order to enable an analysis comparable to the ClinROs:

The daily Peak Pruritus score at the 7 days preceding the visit will be averaged (arithmetic mean) if at least 4 scores are present in the respective time frame. Otherwise, the weekly Peak Pruritus 0-10 NRS will be missing for the respective visit. For end of treatment visits no scores after end of treatment will be considered for the average.

Daily recording of Peak Pruritus 0-10 NRS in the follow-up phase was not possible after Day 23 after end of treatment for many participants due to a technical issue. In this case the recording ended before the planned FU visit at Day 28 after end of treatment. Due to this reason, the weekly Peak Pruritus 0-10 NRS for the FU visit will not be based on the 7 days preceding the FU visit but it will be based on the recording from Day 17 to Day 23 after end of treatment for all participants. Thus, the FU assessment for the weekly Peak Pruritus 0-10 NRS will be a bit earlier than that for the other efficacy assessments.

The following will be considered for imputation procedures: Weekly average Peak Pruritus 0-10 NRS scores will only be used for multiple imputation or be carried forward as described in Section 4.2 and 4.3 if the respective visit was at least 7 days after last rescue medication use

(i.e. no rescue medication use during the last 7 days before the visit when the daily Peak Pruritus scores are assessed which are considered for the weekly average. If this is not possible then an earlier weekly average score will be used for imputation, the baseline weekly average score will always be acceptable.

Binary endpoints

For binary endpoints, the secondary efficacy analysis will use the same approach as for the primary analysis (including sensitivity and supplementary analyses).

- *In detail, the proportion of participants reaching EASI 50 response at Week 12 (Day 84) will be analyzed. An EASI 50 response is defined as achievement of 50% reduction from baseline in the Eczema Area and Severity Index.*
- *Likewise, the proportion of participants reaching EASI 90 response at Week 12 (Day 84) will be analyzed. An EASI 90 response is defined as achievement of 90% reduction from baseline in the Eczema Area and Severity Index.*
- *Furthermore, the proportion of participants with a response with respect to vIGA-AD at Week 12 (Day 84) will be analyzed. A vIGA-AD response refers to achievement of a vIGA-AD of 0 or 1 with at least a 2-grade reduction from baseline (vIGA-AD scale ranges from 0-4).*
- *In addition, the proportion of participants with a response with respect to the average of the Peak Pruritus 0-10 NRS at Week 12 (Day 84) will be analyzed. A response with respect to peak pruritus will be defined as achievement of a ≥4 point improvement (reduction) in the weekly average of the Peak Pruritus 0-10 NRS score from baseline (for participants with Peak Pruritus 0-10 NRS score ≥ 4 at baseline) (Phan et al, 2012).*

The logistic regression model used for imputation (Section 4.2.2) will include the EASI score, vIGA-AD score or the weekly average Peak Pruritus 0-10 NRS at baseline and Day 56 (Visit 6), respectively (if possible).

Continuous endpoints

For continuous endpoints, i.e. percent change from baseline in EASI score, absolute change from baseline in BSA as well as percent change from baseline in the weekly average of the Peak Pruritus 0-10 NRS score, the impact of a potential use of rescue treatment or SoC for treatment of AD on the efficacy will be accounted for as follows:

If rescue medication is used between Day 56 (Visit 6) and Day 84 (Visit 7) (or end of treatment if before Day 84 (Visit 7)) *or if SoC is used (at any time) or study drug is discontinued due to lack of efficacy (at any time), data collected afterwards will be handled as missing.*

Missing data will be imputed using a multiple imputation approach followed by an analysis of covariance model.

Imputation of missing EASI score for Day 84 (Visit 7)

The multiple imputation procedure for continuous endpoins will be similar to the one for binary endpoints but using linear regression instead of logistic regression. For a detailed

technical description and data handling steps of the imputation of missing EASI score please also refer to Section 8.4.1 and 8.4.3. In the following the imputation procedure will be described in general:

The following analyses will be performed on the PPS:

In a first step any values (set to) missing after rescue medication use at or after Day 56 (Visit 6), i.e until Day 84 (Visit 7) or end of treatment if before Day 84 (Visit 7), after use of systemic SoC and after discontinuation of study intervention due to lack of efficacy will be imputed by carrying forward the worst observed post-dose value to impute the value at Day 84 (Visit 7) (refer to Section 8.4.3). SoC medication use after early discontinuation due to other reasons will not trigger the imputation by worst observed post-dose value carried forward for the following time points.

The remaining missing values at Day 84 (Visit 7) will be imputed using a multiple imputation procedure based on the absolute values of the EASI score.

The imputation will be performed by treatment group using the monotone linear regression method from the SAS multiple imputation procedure (PROC MI).

For a participant

- with EASI score at baseline and Day 56 (Visit 6),
- without EASI score at Day 84 (Visit 7),
- without systemic SoC use before Day 56 (Visit 6) EASI score,
- without rescue medication use at Day 56 (Visit 6) EASI score

a linear regression model will be built based on the following participants:

- with EASI score at baseline and Day 56 (Visit 6),
- with EASI score at Day 84 (Visit 7),
- without systemic SoC use before Day 56 (Visit 6) EASI score,
- without rescue medication use at Day 56 (Visit 6) EASI score,
- without rescue medication use (irrespectively of the indication) within the 7 days before the Day 56 (Visit 6) EASI score (to avoid influence of EASI scores directly after rescue medication on the imputation of the value at Day 84 (Visit 7)).

Participants with use of systemic SoC or rescue medication or discontinuation of study intervention due to lack of efficacy after the Day 56 (Visit 6) EASI score and until Day 84 (Visit 7) will have an EASI score (worst post dose observation) at Day 84 (Visit 7) and can be included in the linear regression model if they fulfill the criteria specified above.

The EASI score at Day 84 (Visit 7) will be the dependent variable in the linear regression model and the following variables will be included as independent variables:

- EASI score at baseline,
- EASI score at Day 56 (Visit 6).

If the multiple imputation using linear regression as described above is based on less than 8 participants per treatment group the baseline EASI score and the EASI score at the previous visit will be used as independent variable in the linear regression model as described above. If there will not be sufficient participants to include one of the post-dose visits (Day 28 (Visit 5),

or – if not possible – Day 14 (Visit 4)) the linear regression model will only use the EASI score at baseline as independent variable. Then the model will be based on all participants with EASI score at baseline and EASI score at Day 84 (Visit 7). The EASI at Day 84 (Visit 7) could be imputed using worst post-dose observation carried forward after SoC use, discontinuation of study intervention due to lack of efficacy or use of rescue medication after Day 56 (Visit 6). The variables used in the regression model will be indicated below the output table.

If there is a participant with rescue medication use within the 7 days before the Day 56 (Visit 6) EASI score and missing EASI at Day 84 (Visit 7) which needs to be imputed the same procedure as described above (in case of less than 8 particpants per treatment group) will be applied for that participant.

Imputed values below the minimum (0) or above the maximum (72) possible EASI score will be set to the minimum or maximum score.

Imputation according to the procedure defined above will be performed 50 times. All participants with EASI scores imputed using the linear regression method will be combined with the remaining participants (which will have the same values for each imputed dataset).

Then an analysis of covariance (ANCOVA) with the percent change from baseline in the EASI score at Day 84 (Visit 7) as dependent variable will be applied for each imputed dataset. The model will include treatment group as fixed effect and the EASI baseline value as covariate. The results from the model will be combined by Rubin's rule using SAS MIANALYZE procedure. The treatment difference (including 95% confidence intervals) at Day 84 (Visit 7) will be presented.

As a "worst case" scenario the analysis will be repeated by imputing all missing EASI scores at Day 84 (Visit 7) by carrying forward the worst post-dose EASI score (refer to Section 8.4.3 for details).

Furthermore, the analysis will be repeated by imputing all missing EASI scores at Day 84 (Visit 7) using last observation carried forward (refer to Section 8.4.3 for details).

As for the binary endpoints supplementary analyses with respect to the time point after which rescue medication use leads to setting the subsequent values to missing will be performed as described for the binary endpoints in Section 4.2.4. The analysis will be similar to the main analysis but considering rescue medication use for analysis starting from Day 1 (Visit 3) or not at all. These analyses handle the rescue medication use in a similar manner over time. Thus, a mixed effect repeated measurement model will be applied in addition assuming the values are missing at random; and furthermore multiple imputation based on linear regression including all participants will be performed with subsequent ANCOVA analysis (for details refer to Section 4.4.2 and 8.4.3).

The analyses described above will be repeated on the FAS with exception of the main analysis (see Table 4-2). This approach will not be applied due to potential scarceness of participants without use of rescue medication in the 7 days before EASI scores at early time points before Day 56 (Visit 6). Furthermore, the analysis disregarding the use of rescue medication will not be performed on the FAS.

Table 4-2 summarizes the analyses for the percent change in EASI score at Day 84 (Visit 7).

Table 4-2 Main, sensitivity and supplementary analyses for the secondary endpoint percent change from baseline in EASI at Day 84 (Visit 7)

Analysis set	Analysis type	After "Use of rescue medication"	After "Use of SoC / discontinuation due to lack of efficacy"	After "Discontinuation due to other reasons" or missing values	No. of analysis for continuous endpoint
PPS	Main	Worst post-dose observation carried forward at/after Day 56 (Visit 6)	Worst post-dose observation carried forward	Multiple imputation using linear regression	C_1.1 (Main)
	Sensitivity	Worst post-dose observation carried forward at/after Day 56 (Visit 6)	Worst post-dose observation carried forward	Worst post-dose observation carried forward	C_1.2 (Sensitivity 1.1)
		Last observation carried forward at/after Day 56 (Visit 6)	Last observation carried forward	Last observation carried forward	C_1.3 (Sensitivity 1.2)
	Supplementary	Worst post-dose observation carried forward after Day 1 (Visit 3)	Worst post-dose observation carried forward	Worst post-dose observation carried forward	C_2.1 (Supplementary 1.1)
		Last observation carried forward after	Last observation carried forward	Last observation carried forward	C_2.2 (Supplementary 1.2)

		Day 1 (Visit 3)			
		Values set to missing after Day 1 (Visit 3), no imputation	Values set to missing after SoC, no imputation	No imputation	C_2.3* (Supplementary 1.3)
		Multiple imputation using linear regression after Day 1 (visit 3)	Multiple imputation using linear regression	Multiple imputation using linear regression	C_2.4 (Supplementary 1.4)
		Disregard use of rescue medication for analysis	Worst post-dose observation carried forward	Worst post-dose observation carried forward	C_3.1 (Supplementary 2.1)
			Last observation carried forward	Last observation carried forward	C_3.2 (Supplementary 2.2)
			Values set to missing after SoC, no imputation	No imputation	C_3.3* (Supplementary 2.3)
			Multiple imputation using linear regression	Multiple imputation using linear regression	C_3.4 (Supplementary 2.4)
FAS	Supplementary	Worst post-dose observation carried forward at/after Day 56 (Visit 6)	Worst post-dose observation carried forward	Worst post-dose observation carried forward	C_4.1 (Supplementary 3.1)

		Last observation carried forward at/after Day 56 (Visit 6)	Last observation carried forward	Last observation carried forward	C_4.2 (Supplementary 3.2)
		Worst post-dose observation carried forward after Day 1 (Visit 3)	Worst post-dose observation carried forward	Worst post-dose observation carried forward	C_5.1 (Supplementary 4.1)
		Last observation carried forward after Day 1 (Visit 3)	Last observation carried forward	Last observation carried forward	C_5.2 (Supplementary 4.2)
		Values set to missing after Day 1 (Visit 3), no imputation	Values set to missing after SoC, no imputation	No imputation	C_5.3* (Supplementary 4.3)
		Multiple imputation using linear regression after Day 1 (visit 3)	Multiple imputation using linear regression	Multiple imputation using linear regression	C_5.4 (Supplementary 4.4)

* Analyses will be performed with a mixed effect repeated measurement model as described in Section 4.4.2. All other analyses will be performed with an ANCOVA model as described above.

For the other continuous endpoints, i. e. absolute change of BSA affected by AD and percent change of the weekly average of the Peak Pruritus 0-10 NRS score from baseline at Day 84 (Visit 7) the same analyses as described above will be performed.

After applying multiple imputation using linear regression the imputed values will be restricted to the following ranges by setting values below the minimum to the minimum and values above the maximum to the maximum possible value:

- BSA affected by AD: 0 – 100 %,
- Peak Pruritus 0-10 NRS: 0 – 10.

4.4 Other Endpoint(s) Analysis

4.4.1 Achievement of EASI 75 response up to Week 8 (Day 56)

Generally, the EASI 75 response over time will be analyzed similarly to the analyses for Day 84 (Visit 7) but restricted to those approaches where the intercurrent event "use of rescue medication" is handled similary between Day 1 (Visit 3) and Day 84 (Visit 7) to have comparable estimates over time. The analysis will be done according to the exploratory estimand for all time points (refer to Section 1.1). In detail, the following analyses will be performed for all time points (refer to Table 4-1):

- PPS, B_2.1, B_2.2 ("use of rescue medication" is handled as non-response for all following time points),
- FAS, B_7.1 and B_7.2 ("use of rescue medication" is handled as non-response for all following time points).

In addition, an analysis over time will be performed on the PPS where rescue medication does not trigger the classification as non-responder at all to assess the effect of study intervention on top of topical AD medication:

- PPS, B_5.1 and B_5.2 ("use of rescue medication" is ignored for all time points).

For details with respect to data handling and the imputation strategy refer to Section 8.4.

The probabilities for EASI 75 response at Day 84 (Visit 7) by treatment and the difference between treatments as derived from the Bayesian analysis will be plotted including 95% credible intervals for all analyses as summarized in Table 4-1.

Furthermore, the probabilities for EASI 75 response over time by treatment and the difference between treatments as derived from the Bayesian analysis as described above will be plotted including 95% credible intervals for the following analyses in one plot (for numbers refer to Table 4-1):

- PPS, B_2.1 + B_2.2 and B_5.1 + B_5.2,
- FAS, B_7.1 + B_7.2.

The other binary efficacy endpoints will be analyzed similarly.

4.4.2 Percent change in EASI score over time

As for binary endpoints, the percent change in EASI score over time will be analyzed similarly to the analyses for Day 84 (Visit 7) but restricted to those approaches where the intercurrent event "use of rescue medication" is handled similary between Day 1 (Visit 3) and Day 84 (Visit 7) to have comparable estimates over time. In detail, the following analyses will be performed for all time points (refer to Table 4-2):

- PPS, C_2.1 and C_2.2 (values after first “use of rescue medication” are set to missing and imputed for the following time points),
- FAS, C_5.1 and C_5.2 (values after first “use of rescue medication” are set to missing and imputed for the following time points).

In addition, an analysis over time will be performed on the PPS where use of rescue medication is disregarded for analysis to assess the effect of study intervention on top of topical AD medication:

- PPS, C_3.1 and C_3.2 (“use of rescue medication” is ignored for all time points).

Furthermore, for both scenarios as described above, i. e.:

1. values after first “use of rescue medication” are set to missing for the following time points
   (PPS: analysis C_2.3, FAS: analysis C_5.3),
2. “Use of rescue medication” is ignored for all time points
   (PPS: analysis C_3.3)

a *mixed-effect repeated measurement model will be applied. This model will include the respective continuous efficacy endpoint as dependent variable. All post-dose data* (without FU visit) *will be included in the model. The model will include the respective baseline value as covariate, treatment group, visit and the interaction of visit and treatment group (visit*treatment group) as fixed effects. The covariance structure will be assumed to be unstructured, but the same for both treatment groups. The structure of the covariance matrix might be adapted, e. g. in case of convergence problems. The treatment difference (including 95% confidence intervals) at the planned end of treatment at Day 84 (Visit 7), as well as at the other time points, will be calculated by appropriate contrasts from the mixed-effect repeated measurement model.*

Single subjects without any valid post-dose value in the treatment phase according to Section 8.4.1 (e. g. which discontinued treatment before first post-dose visit at Day 14 (Visit 4) due to lack of efficacy or took rescue medication at all visits until treatment discontinuation) will be included in the mixed model by using last observation carried forward for Day 14 (Visit 4) to avoid bias. Post-dose values from unscheduled visits or early discontinuation visits will be considered irrespectively of any time windows (but early discontinuation visits will only be considered if study intervention was not stopped earlier than 3 days ago). Otherwise, the baseline value will be carried forward.

Additionally for the same scenarios as for the mixed effect repeated measurement model, multiple imputation using linear regression will be performed using the multiple imputation procedure (PROC MI) from SAS based on the following variables: treatment, absolute EASI score at Day 1 (Visit 3), Day 14 (Visit 4), Day 28 (Visit 5), Day 56 (Visit 6), Day 84 (Visit 7). The fully conditional specification (FCS) method will be used in such a way that for each imputed continuous variable, all preceding variables will be considered. The seed will be 6345 and 50 imputations will be performed. Imputed values will be restricted to the range of possible values using the minimum and maximum options of SAS multiple imputation procedure (i. e. if values are imputed outside of the possible range, another value will be re-

drawn for imputation). The missing data pattern and variance information outputs for the multiple imputation will be presented in Section 10 of the CSR.

All participants from the PPS or FAS, respectively, will be included.

If the linear regression will not be possible for both treatment groups due to scarcity of data especially at later time points the imputation will only be performed for the zabedosertib group.

Subsequently, an analysis of covariance (ANCOVA) as described above will be performed by visit for all visits as planned.

The mean estimates for percent change of EASI score at Day 84 (Visit 7) by treatment and the difference between treatments as derived from the ANCOVA model or mixed model of repeated measures analysis as described above will be plotted including 95% confidence intervals for all analyses as summarized in Table 4-2.

Furthermore, the mean estimates for percent change in EASI score over time by treatment and the difference between treatments as derived from the ANCOVA model or mixed model of repeated measures analysis as described above will be plotted including 95% confidence intervals for the following analyses in one plot (for numbers refer to Table 4-2):

- PPS, C_2.1 + C_2.2 + C_2.3 + C_2.4 and C_3.1 + C_3.2 + C_3.3 + C_3.4,
- FAS, C_5.1 + C_5.2 + C_5.3 + C_5.4.

Other continuous end points will be handled similarly.

4.4.3 Further binary endpoints

The other pre-specified binary endpoints, i. e. achievement of EASI 100 response and achievement of a ≥ 3 point improvement (reduction) in the weekly average of the Peak Pruritus 0-10 NRS from baseline at Week 12 (Day 84) will be analyzed similar to the primary endpoint.

4.4.4 Descriptive statistics for binary efficacy endpoints for all time points

The number and frequency of participants with EASI 75 response will be summarized

- for each scenario described in Section 4.2 (refer to Table 4-1) (denominator = number of all participants in analysis set due to imputation of all missing values),
- for the raw data including all values as recorded (denominator = number of participants in analysis set per visit),
- for raw data after setting all values after use of rescue medication at Day 56 (Visit 6) or later and after use of SoC to missing (denominator = number of participants in analysis set per visit),
- for raw data after setting all values after use of rescue medication at Day 1 (Visit 3) or later and after use of SoC to missing (denominator = number of participants in analysis set per visit).

If EASI 75 response will be imputed using logistic regression, several (potentially different) EASI 75 responses will be available per participant. In this case the responses of the participant will be averaged over all imputed data sets. The average will be used to count the number and frequency of participants with EASI 75 response (i. e. non-integer numbers of EASI 75 responders are possible to summarize frequencies most accurately).

All other binary endpoints will be summarized similarly.

In addition, the number and frequency of participants with vIGA-AD scores (0, 1, 2, 3, 4) will be summarized (the three summaries based on raw data as described above, no imputation will be performed for vIGA-AD values).

4.4.5 Descriptive statistics for continuous efficacy endpoints for all time points

Summary statistics will be presented for EASI scores including percent change and absolute change from baseline.

- for imputed data after each type of imputing as outlined above (refer to table in Section 4.3.2),
- for raw data including all values as recorded,
- for raw data after setting all values after use of rescue medication at Day 56 (Visit 6) or later and after use of SoC to missing,
- for raw data after setting all values after use of rescue medication at Day 1 (Visit 3) or later and after use of SoC to missing.

If EASI scores will be imputed using linear regression, several different EASI scores will be available per participant. In this case the scores of the participant will be averaged over all imputed data sets. The average will be used for calculation of summary statistics.

The summary statistics for raw data will also comprise changes from end of treatment to follow-up.

For imputed data after each type of imputing (refer to table in Section 4.3.2) the following outputs will be presented as well:

Histograms will be presented together with a box-plot for percent change from baseline in EASI score at Day 84 (Visit 7) by treatment group (in one plot).

In addition, a frequency table will be presented for the main analysis summarizing percent change in EASI from baseline at Day 84 (Visit 7) by -100% to -90%, >-90 to -80%, >-80 to -70%, … , >-20 to 10%, >-10 to <0%, 0 to 10%, >10% to 20% etc. including cumulative frequencies.

Individual plots of EASI scores (raw data including all values as measured) will be presented as well based on the FAS. EASI scores after systemic SoC use or use of rescue medication at or after Day 56 (Visit 6) will be marked.

Similar outputs will be presented for weekly average of Peak Pruritus 0-10 NRS and BSA affected by AD (BSA with absolute changes and change categories of 5% steps for frequencies).

In addition, mean plots over time of daily Peak Pruritus 0-10 NRS score will be presented (similar to the three summaries based on raw data as described above, no imputation will be performed for daily peak pruritus values).

Relations between endpoints

Relations between outcome variables will be analyzed on the PPS after imputing according to model C_1.1 + B_1.1 (refer to Section 4.3.2 and 4.2.4). To examine the relationship between the response variables with respect to treatment effect at Day 84 (Visit 7) a spiderweb will be presented by treatment group displaying the following values on the axes of the web:

- mean percent change in EASI score from baseline,
- mean percent change in weekly average of Peak Pruritus 0-10 NRS from baseline,
- mean absolute change in BSA affected by AD from baseline,
- frequency of achievement of a vIGA-AD response (score 0 or 1 and ≥ 2 points improvement),
- frequency of achievement of 75% reduction from baseline in the Eczema Area and Severity Index (EASI 75 response),
- frequency of achievement of a ≥ 4 point-improvement (reduction) in the weekly average of the Peak Pruritus 0-10 NRS score from baseline.

The mean values will be normalized with respect to 99th percentile minus 1st percentile. The frequencies will be presented on an axis from 0 to 1. The mean changes and frequencies at Day 84 (Visit 7) will be connected by lines. The points on the axes reflecting no change (zero) or a frequency of 0% will be presented and connected by lines as well.

Scatter plots including Pearson and Spearman correlation coefficients will be presented on the PPS after imputing according to model C_1.1 for:

- absolute change in EASI at Day 84 (Visit 7) from baseline versus absolute change in BSA affected by AD at Day 84 (Visit 7) from baseline,
- absolute change in EASI at Day 84 (Visit 7) from baseline versus absolute change in weekly average of Peak Pruritus 0-10 NRS at Day 84 (Visit 7) from baseline,
- absolute change in BSA affected by AD at Day 84 (Visit 7) from baseline versus absolute change in weekly average of Peak Pruritus 0-10 NRS at Day 84 (Visit 7) from baseline.

The three scatter plots above will also be presented showing percent change from baseline for both respective variables.

Furthermore, individual plots will be presented on the SAF with the time course of absolute EASI score, absolute weekly average of Peak Pruritus 0-10 NRS, absolute BSA affected by AD, and vIGA-AD in one plot per participant. Raw data will be presented. Use of systemic SoC and use of rescue medication (at any time) will be presented as boxes from start to end of use. Date of discontinuation of study intervention will also be indicated (in case of early

discontinuation of study intervention differing by early discontinuation due to lack of efficacy or other reasons). The study day will be presented on the x-axis. Participant identifier and validity for analysis sets will be indicated above each plot as PPS (if in PPS), FAS (if in FAS and not PPS) or SAF (if not in PPS and not in FAS).

Influence of baseline characteristics

The influence of baseline characteristics will be analyzed on the PPS after imputing according to model C_1.1 (see Section 4.3.2, the scores derived from multiple imputation for each participant will be averaged over all imputed data sets for this analysis). To identify the most influential set of baseline characteristics explaining the change in the EASI score at Day 84 (Visit 7) regression analyses will be performed with percent change in EASI score from baseline at Day 84 (Visit 7) as dependent variable and treatment and the following baseline characteristics as predictors:

- treatment (zabedosertib or placebo),
- age,
- body mass index (BMI) at baseline,
- sex,
- race (only if at least two categories with at least n=5 per treatment group),
- EASI at baseline,
- BSA affected by AD at baseline,
- vIGA-AD at baseline,
- weekly average of Peak Pruritus 0-10 NRS at baseline,
- disease duration at baseline,
- age of onset,
- number of AD flares in the year before informed consent,
- history of asthma, allergic rhinitis or allergic conjunctivitis (yes/no),
- IgE level at baseline
- previous treatment with systemic standard of care for AD within 6 months (i. e. 183 days) before start of study drug (yes/no)
  (refer to Section 6.3 for a list with respective medication).

Univariate regression with only one variable per regression model will be performed and a multiple regression model will be applied using ‘forward selection’. This stepwise procedure starts with no explanatory variables in the model, then testing the explanatory variables one by one in univariate models. The most significant variable will be added to the model given the p-value is below a certain threshold (≤0.2). Then all variables will be added one by one to the current model and the most significant one will be added if the p-value is ≤0.2. This step will be iterated until no further variables can be added to the model, i. e. p>0.2 for all remaining variables.

4.4.6 Descriptive statistics for main intercurrent events

The number of participants with intercurrent event (ICE as defined in Section 8.3) 'use of systemic standard of care', with 'use of rescue medication', with 'discontinuation of study intervention due to lack of efficacy' and with 'discontinuation of study intervention due to other reasons' (i.e. not due to lack of efficacy) will be presented by treatment group and visit interval including cumulative numbers. Percentages will be calculated with respect to the number of participants in the analysis set.

The probability of the following events to occur during the treatment phase will be summarized for weekly intervals using the life table method:

- time to first use of rescue medication at or after Day 56 (Visit 6), first use of systemic SoC or discontinuation of study intervention due to lack of efficacy (all ICE as defined in Section 8.3), whatever comes first,
- time to first use of rescue medication at or after Day 1 (Visit 3), first use of systemic SoC or discontinuation of study intervention due to lack of efficacy (all ICE as defined in Section 8.3), whatever comes first,
- time to first use of systemic SoC or discontinuation of study intervention due to lack of efficacy (all ICE as defined in Section 8.3), whatever comes first.

The probability of an event to occur for the first time in the interval x_i to x_{i+1} will be calculated as follows:

Let

- $n_{x(i)}$ be the number of participants who had no such event until time x_i,
- $r_{x(i)}$ be the number of participants with the first occurrence of such an event in the interval x_i to x_{i+1},
- $c_{x(i)}$ be the number of participants censored in the interval x_i to x_{i+1}
  Censoring will occur after early discontinuation of treatment due to other reasons (i.e. not lack of efficacy), after regular treatment completion and after use of rescue medication after the respective time point or use of SoC other than the intercurrent events defined in Section 8.3 (i.e. use not related to AD),
- $n'_{x(i)} = n_{x(i)} - \frac{1}{2} c_{x(i)}$ be the effetive number of participants at risk for first occurrence of such an event in the interval x_i to x_{i+1}.

Then

- $p_{x(i)} = \frac{r_{x(i)}}{n'_{x(i)}}$ is the probability for first occurrence of such an event in the interval x_i to x_{i+1} given that no event occurred until time x_i,
- $P_{x(i)} = 1 - (1 - p_{x(0)})(1 - p_{x(1)}) \cdots (1 - p_{x(i)})$ is the probability for first occurrence of such an eventuntil time x_{i+1}.

Table 4-3 Example for calculation of life table estimates

Time interval x to $x + \Delta x$ [weeks]	$n_{x(i)}$	$r_{x(i)}$	$c_{x(i)}$	$n'_{x(i)}$	$1 - p_{x(i)}$	$p_{x(i)}$	$1 - P_{x(i)}$	$P_{x(i)}$
Zabedosertib								
0-1	100	2	0	100	0.98 (1-2/100)	0.02 (2/100)	0.98	0.020
1-2	98	3	1	97.5	0.969 (1-3/97.5)	0.031 (3/97.5)	0.949	0.050
2-3	94	5	2	93	0.946 (1-5/93)	0.054 (5/93)	0.898	0.101
3-4	87	3	0	87	0.966 (1-3/87)	0.034 (3/87)	0.867	0.133
…								
11-12								
Placebo								
…								

Note: Columns with gray text are for illustration purpose only and will not be presented in the outputs.

Furthermore, Kaplan-Meier curves for time of first occurrence of such an event will be presented by treatment group.

In addition, a graphical listing will be presented visualizing the time points of the visits and the intercurrent events as defined in Section 8.3 (use of rescue medication, use of SoC, early discontinuation of study intervention due to lack of efficacy) and the intercurrent event early discontinuation of study intervention due to other reasons not related to AD (not due to any ICE that is defined in Section 8.3) for each participant.

4.4.7 Pharmacokinetics

All PK analyses will be performed on the PKS.

Plasma samples will be collected for measurement of plasma concentrations of zabedosertib and its metabolite BAY 2822815 at Day 14 (Visit 4), Day 28 (Visit 5), Day 56 (Visit 6), Day

84 (Visit 7) pre-dose as well as between 0.5 to 1.5 hours and between 2 to 5 hours post-dose at Day 1 (Visit 3) and Day 84 (Visit 7).

T*he following statistics will be reported for plasma concentrations of zabedosertib* and its metabolite *at each visit (and time point if applicable): geometric mean, geometric standard deviation (re-transformed standard deviation of the logarithms) and geometric coefficient of variation (CV), arithmetic mean, standard deviation and CV, minimum, median, maximum value and the number of measurements as well as the number of measurements ≥ limit of quantification (LLOQ). Means at any visit will only be calculated if at least 2/3 of the individual data were measured and were above LLOQ.* Otherwise only minimum, median and maximum will be presented. *For the calculation of the mean value*, standard deviation and median *a data point below LLOQ will be substituted by one half of this limit. In tables showing mean values, where values below LLOQ are included in the calculation of mean values, these means will be marked.*

Furthermore, geometric means including one standard deviation range will be plotted for zabedosertib and its metabolite. The post-dose time between 0.5 to 1.5 hours post-dose will be presented at 1hour post-dose and the time between 2 to 5 hours will be presented at 3.5 hours post-dose on the x-axis.

Box plots including geometric mean for the concentration of zabedosertib and its metabolite BAY 2822815 will be presented as well.

Furthermore, individual plots of zabedosertib and its metabolite will be presented.

Pre-dose PK concentrations will be considered valid for inclusion into summary statistics if there were 8-16 hours between the last documented tablet intake before pre-dose PK sample collection and the pre-dose PK sample collection. Post-dose PK concentrations will be considered valid for inclusion into summary statistics if a tablet intake in the morning of the post-dose PK sample collection was documented. Furthermore, pharmacokinetic values might be flagged as invalid by the pharmacokinetic experts to be excluded from the calculation of summary statistics if applicable.

Zabedosertib is known to bind to alpha glycoprotein in plasma. Scatterplots of absolute concentration of alpha glycoprotein versus zabedosertib concentration in plasma (pre-dose) will be presented for Day 14 (Visit 4), Day 28 (Visit 5), Day 56 (Visit 6), Day 84 (Visit 7) in one plot (different symbols for the different visits will be used). Pearson and Spearman correlation coefficients will be given for the correlation over all visits combined.

4.4.8 Biomarkers

Post-dose biomarker samples will be considered as valid if no topical or systemic AD medication as defined in Section 8.3 (irrespectively of the indication) has been used within the last 7 days before the day of biomarker sample collection and if study intervention was not stopped earlier than 3 days before the day of biomarker sample collection.

Skin biopsy

Skin biopsy samples will be collected at Day 1 (Visit 3) and Day 84 (Visit 7, EOT) from a subgroup of participants willing to participate in this substudy. The analysis will be performed

with participants in the biomarker analysis set who participated in the biopsy substudy and have a biopsy at both visits.

The following assessments with the skin biopsy are optional (the list could be subject to change):

- Skin thickness (stratum corneum to stratum basale)
- Ki67 or MKI67 expression (no. of positively stained cells)
- CK16 or KRT16 expression (IRS score)

Summary statistics including absolute and percent changes to baseline will be presented for each assessment. A box plot will be presented showing the baseline and EOT assessments as well as the absolute change from baseline. Individual values will be plotted as well.

Biomarkers in blood

Blood samples for the inflammatory biomarker analysis will be collected at Day 1 (Visit 3) and Day 84 (Visit 7, EOT).

The analyses will be performed on the biomarker analysis set. Only participants with a blood sample at both visits will be considered.

The following inflammatory biomarkers will be analyzed (the list could be subject to change):

- TSLP (thymic stromal lymphopoietin),
- IL31 (Interleukin 31),
- IL33 (Interleukin 33),
- IL4 (Interleukin 4),
- IL13 (Interleukin 13),
- IL22 (Interleukin 22),
- CCL17 (C-C motif chemokine ligand 17).

Summary statistics including absolute changes to baseline will be presented for each inflammatory marker. Furthermore, summary statistics will also be presented for ratios to baseline using summary statistics as intended for log-normally distributed variables (geometric mean, SD and CV, minimum, median, maximum). For handling of inflammatory markers below the lower limit or above the upper limit of quantification refer to Section 4.1.1.

A box plot will be presented showing the baseline and EOT assessments as well as the ratio to baseline. Individual values will be plotted as well.

For each biomarker, an explorative ANCOVA will be performed with the log-transformed ratio-to-baseline at Day 84 (Visit 7) as dependent variable. The model will include treatment group as fixed effect and the log-transformed baseline concentration as covariate. Least square means will be calculated for each treatment group and two-sided 95% confidence intervals based on the t-distribution and the residual error of the model will be computed for the difference between treatment groups and back-transformed to the original scale.

4.5 Safety Analyses

All safety analyses will be performed on the SAF. Participants will be analyzed according to the treatment they actually received.

4.5.1 Extent of Exposure

Extent of exposure tables will be presented for the SAF, PPS and FAS.

Treatment duration will be defined as the number of days from the day of first study intervention intake up to and including the day of last study intervention intake excluding any gaps in study intervention intake (according to exposure data from eCRF) and will be summarized using descriptive statistics. In addition, the treatment duration will be categorized into weekly intervals (1-7 days, 8-14 days, 15-21 days, … ,78-84 days, > 84 days) and the frequency of participants per interval will be summarized.

The total treatment dose (according to drug accountability data from eCRF) will be summarized in milligram using descriptive statistics. In addition, the treatment dose will be categorized into groups of 2000 mg (0 – 2000 mg, >2000 – 4000 mg etc.) and the frequency of participants per group will be summarized.

The treatment compliance will be calculated between start and end of treatment as the number of tablets actually taken divided by the number of planned tablets multiplied by 100 with:

- Number of tablets actually taken (based on drug accountability data):
  no. of tablets dispensed – no. of tablets returned – no. of tablets lost

  In case of unreturned study drug intervention bottles study drug intake documentation from the hand-held device on daily basis will be used to replace the missing drug accountability data unless further information is provided by the participant (and documented via number of lost tablets). Otherwise, if no further information is available and no replacement is possible from diary data it will be assumed that no tablets have been taken.

- Number of planned tablets:
  The number of planned tablets will be calculated based on the planned dose of 2 tablets per day for each day starting with first intake of study intervention to the last intake of study intervention. For the first and last day of study intervention intake only one planned tablet will be considered if the time of first intake is after 3 pm or the time of last intake is before 3 pm (if no time is available the default will be two doses on the first day of study intervention intake and one dose on the last day). The compliance will be summarized descriptively. In addition, the compliance will be categorized into three groups of <80%, 80-120%, and >120% and the frequency of participants per group will be summarized.

In addition, the following compliance calculations will be needed to check validity for the PPS: treatment compliance will be calculated and summarized presenting the percent of tablets actually taken with respect to the planned number of tablets

a) between start of treatment and end of treatment or start of rescue medication (irrespectively of indication) at or after Day 56 (Visit 6), whatever comes first,

b) during the last 4 weeks before end of treatment or start of rescue medication (irrespectively of indication) at or after Day 56 (Visit 6), whatever comes first.

Compliance calculations as follows will be needed to check for validity for the BAS: treatment compliance will be calculated and listed presenting the percent of tablets actually taken with respect to the planned number of tablets

c) during the last 4 weeks before end of treatment.

Study intervention intake documented by the participant on the hand-held device on daily basis will be used to calculate the number of tablets actually taken for a)-c). If the documentation in the diary has gaps (e. g. due to technical reasons or if the end date documented in the eCRF is more than one day after the last study drug intake documented in the diary) this information will be replaced by drug accountability data, if possible. If data from the diary is replaced by drug accountability data complete intervals of drug accountability data will be used, i. e. between dispense and return date of the medication or treatment end date if medication returned later (in case of bottles returned after dispensation of the next kit overlapping drug accountability data will be combined). If drug accountability intervals don't fit the required interval diary data will be considered to subtract the tablets taken on the extra days if possible (i. e. if the respective diary records are available). If not possible the diary data will be used as is.

In case of unresolved discrepancies between study intervention intake documented in the participant hand-held device and the study intervention intake assessed by counting returned tablets it will be discussed during blind data review if the study intervention intake documentation is considered reliable.

Frequencies of participants with study intervention interruptions will be presented overall and by reason for the interruption and duration of study intervention interruption.

4.5.2 Adverse Events

The secondary safety endpoint is the frequency and severity of TEAEs.

Any AEs related to study procedures recorded after signing of informed consent but prior to first treatment with study intervention will be considered as pre-treatment AEs. *TEAEs are events that start after the first treatment with the study intervention and until 7 days after the last intake of study intervention.* AEs that occurred 8 days after last dose or later will be considered as post-treatment AEs.

All TEAEs will be listed.

Individual listings of all AEs will be provided indicating whether the AE is treatment emergent and including AE onset and end date relative to start and end date of treatment and the duration of the AE.

The following AEs have been defined as AEs of special interest and are recorded as such:

- Confirmed or suspected severe invasive bacterial infections,
- Systemic hypersensitivity reactions,
- Non-invasive infections (skin),
- Unexplained cases of rhabdomyolysis.

Adverse events will be coded using medical dictionary for regulatory activities (MedDRA) terms using the most recent effective MedDRA version as per TS domain.

Overall summaries of the number of participants with any AEs, pre-treatment AEs, TEAEs, and post-treatment AEs will be presented. The overall summary will comprise the frequency of participants with certain AEs: Participants with

- any AE,
- any AE by maximum intensity (mild, moderate, severe),
- any AE related to study drug,
- any AE related to study drug by maximum intensity (mild, moderate, severe),
- any AE related to study procedures,
- any AE leading to discontinuation of study drug,
- any AE of special interest (also broken down by category of special interest as stated above),
- any SAE,
- any SAE related to study drug (maximum intensity),
- any SAE related to study procedures,
- any SAE leading to discontinuation of study drug,
- any AE with outcome death.

The frequency of participants with

- SAEs
- TEAEs,
- TEAEs by maximum intensity,
- TEAEs related to study drug,
- TEAEs related to study drug by maximum intensity,
- TEAE of special interest,
- TEAE of special interest by maximum intensity,
- TEAEs of special interest related to study drug,
- TEAEs resulting in discontinuation of study drug,
- TEAEs resulting in discontinuation of study drug by maximum intensity,
- TEAEs resulting in discontinuation of study drug and related to study drug,
- serious TEAEs,
- pre-treatment AEs,
- post-treatment AEs

will be summarized using the MedDRA terms system organ class (SOC) and preferred term (PT).

In case of events with different intensity within a participant, the maximum reported intensity will be used. If intensity is missing, the event will be considered as severe. If the same event is reported as both unrelated and related to the study intervention within a participant, the event will be considered as related to study intervention. If the study intervention relationship is missing, the event will be considered as being related to the study intervention.

Death, TESAEs, TEAESIs or TEAEs leading to discontinuation of study intervention, TEAEs not resolved at end of study *will be listed separately.*

Additionally, a listing of participants with COVID-19 adverse event will be presented (with COVID-19 AEs only and another listing with all AEs of participants with COVID-19 AE).

4.5.3 Additional Safety Assessments (if applicable)

Laboratory assessments

Continuous *laboratory values will be summarized descriptively by parameter including changes from baseline. Arithmetic mean plots including standard deviation over time will be presented* (LLN and ULN will be indicated if applicable, in case of multiple normal ranges, the lowest LLN and highest ULN will be presented). For laboratory parameters with multiple normal ranges the summary statistics will also be broken down by normal range. In addition, the lab values will be normalized with respect to the respective range and an additional summary statistic will be presented for the normalized values. Normalization will be performed as follows: (value – LLN)/(ULN – LLN). In addition, box plots including arithmetic means over time will be presented for changes from baseline. Individual plots of laboratory values will also be presented (in case of multiple normal ranges separate plots for each normal range).

For the laboratory parameters aspartate aminotransferase (AST/SGOT), alanine aminotransferase (ALT/SGPT), creatine kinase, high sensitivity C reactive protein (hsCRP) concentration, as well as white blood cell, neutrophil and lymphocyte counts, a proc mixed repeated measurement model will be calculated with the change from baseline at all post-dose visits as dependent variable. The model will include treatment group, visit and the interaction of visit and treatment group (visit*treatment group) as fixed effects and the respective baseline concentration as covariate. The covariance structure for the within-subject correlation between outcomes at different timepoints will be assumed to be unstructured. The structure of the covariance matrix might be adapted, e. g. in case of convergence problems. Least square means will be calculated for each treatment group and two-sided 95% confidence intervals based on the t-distribution and the residual error of the model will be computed for the difference in treatment means.

The number of participants with treatment-emergent high (>ULN) or low (<LLN) abnormal laboratory values will be summarized for each laboratory parameter.

In addition, hepatocellular drug-induced liver injury screening plots will be provided as decribed in the FDA standard safety tables and figures: integrated guide document. The hepatocellular drug-induced liver injury screening plot will present each participant based on their maximum post-baseline total bilirubin (y-axis) and transaminase (ALT or AST, whichever is higher). Each value is expressed as multiples of the ULN. Dashed lines in this

plot represent total bilirubin and transaminase cutoffs of 2 x ULN and 3 x ULN, respectively and are based on Hy's Law criteria. Potential Hy's Law cases will be circled red and are defined as having any post-baseline total bilirubin equal to or exceeding 2 x ULN within 30 days after a post-baseline ALT or AST equal to or exceeding 3 x ULN, and maximum post-baseline ALP <2 x ULN. Furthermore, a cholestatic drug-induced liver injury screening plot will present each participant by their maximum ALP versus their maximum total bilirubin values in the post-baseline period. A potential cholestatic drug-induced liver injury case (red circled) will be circled red and is defined as having a maximum post-baseline total bilirubin equal to or exceeding 2 x ULN within 30 days after post-baseline ALP became equal to or exceeding 2 x ULN.*For laboratory values shift tables will be provided (by time point including any time point* during treatment phase*) focusing on the frequency and percentage of participants with shifts from within normal range (or below normal range) at baseline to above normal range post-baseline and from within normal range (or above normal range) at baseline to below normal range post-baseline.*

Frequency tables will be provided for semi-quantitative and qualitative data.

Courses of abnormal laboratory values (continuous and qualitative data) at least at one visit will be listed by parameter as absolute values and – if applicable – as multiples of the respective upper or lower limit (abnormal values will be flagged with 'H' for high and 'L' for low).

Furthermore, a frequency table with the number of subjects with the following treatment-emergent laboratory abnormalities will be presented. A laboratory abnormality is considered to be treatment-emergent if it occurs after treatment start and up to 7 days after end of treatment and has not been present before treatment start. The following abnormalities will be considered:

Laboratory abnormalities leading to study drug discontinuation:

- WBC count < 1000 cells/µL; < 1.0 x 10^9 /L,
- Absolute neutrophil count < 500 cells/µL; < 0.5 x 10^9 /L,
- Lymphocyte count < 200 cells/µL; < 0.2 x 10^9 /L,
- Hemoglobin < 8 g/dL,
- Creatine kinase (CK) > 10 x ULN

Liver abnormalities:

- ALT or AST > 3 x ULN
- ALT or AST > 3 x ULN and total bilirubin (TBL) > 2 x ULN
- ALT or AST > 3 x ULN and INR > 1.5 x ULN
  (INR > 1.5 x ULN is only relevant if the participant is not on Vitamin K antagonist, new oral anticoagulants (NOACs) or heparin.)
- ALT or AST > 5 x ULN for more than 2 weeks
- ALT or AST > 8 x ULN

Abnormalities with respect to neutrophil counts:

- Absolute neutrophil count < 500 cells/µL,
- Absolute neutrophil count 500 to < 1000 cells/µL,
- Absolute neutrophil count 1000 to < 1500 cells/µL.

Abnormalities with respect to creatinine kinase:

- Creatine kinase (CK) >ULN to < 5 x ULN,
- Creatine kinase (CK) 5 to 10 x ULN,
- Creatine kinase (CK) > 10 x ULN.

Vital signs

Vital signs will consist of single measurements of systolic blood pressure, diastolic blood pressure, pulse rate and body temperature. Vital signs including height (only summary statistics for the baseline value), weight and BMI will be summarized descriptively by parameter including absolute changes from baseline. Arithmetic mean plots including standard deviation over time will be presented. In addition, box plots including arithmetic means over time will be presented for changes from baseline. Individual plots over time will be generated.

Electrocardiogram

Single 12-lead ECG will be obtained locally.

ECG parameters (PR, QRS, QT, corrected QT intervals (according to Bazett's formula) and the ventricular rate) will be summarized descriptively by parameter including absolute changes from baseline. Arithmetic mean plots including standard deviation over time will be presented. In addition, box plots including arithmetic means over time will be presented for changes from baseline. Individual plots over time will be presented.

Furthermore, frequency tables will be presented for corrected QT intervals for absolute values with the following categories:

- >450 to 480 msec, >480 msec to 500 ms, >500 msec,

and absolute change from baseline with the following categories:

- >30 to 60 msec, >60 msec.

A frequency table will be provided for the ECG findings assessed by the investigator, and for overall interpretation of ECGs (normal; abnormal, clinically insignificant; abnormal, clinically significant).

4.6 Other Analyses

Not applicable.

4.6.1 Other Variables and/or Parameters

Not applicable.

4.6.2 Subgroup Analysis

The following subgroups will be defined (ATC codes to identify the respective previous systemic treatments are given in brackets):

- EASI score at baseline (≤/> median),
- vIGA-AD at baseline (=3 / =4),
- BSA affected by AD at baseline (≤/> median),
- Weekly average of the Peak Pruritus 0-10 numerical rating scale (NRS) score at baseline (≤ / > median),
- IgE level at baseline (<300 kU/L, 300-1000 kU/L, > 1000 kU/L),
- age ≥18 – <40, ≥40 - <= 65,
- age of onset of AD (<18 / >=18 years),
- sex,
- race,
- body mass index (BMI) at baseline (<25, 25 - 30, >30),
- previous treatment with systemic standard of care for AD within 6 months (i.e. 183 days) before start of study drug (yes/no)
  (refer to Section 6.3 for a list with respective medication).
- previous treatment with systemic glucocorticoids/ non-steroidal systemic immunosuppressants/ calcineurin inhibitors for AD (Yes/No)
  (- H02AB/H02BX Glucocorticoids for systemic use
  - L04AA06 MMF (Mycophenolic acid, Mycophenolate sodium, Mycophenolate acid, Mycophenolate mofetil)
  - L04AX01 Purine antagonist – Azathioprine
  - L04AX03, L01BA01 Antineoplastic folic acid analogues – Methotrexate
  - L04AD 01 Ciclosporin
  - L04AD 02 Tacrolimus
  - L04AD 03 Voclosporin),
- previous treatment with Dupilumab/Tralokinumab for AD (Yes/No)
  (- D11AH05, IL-4 and IL-13 monoclonal antibody – Dupilumab
  - D11AH07 IL-13 antibodies – Tralokinumab),
- previous treatment with JAK inhibitor for AD (Yes/No)
  (- L04AA37 JAK-inhibitor – Baricitinib
  - L04AA44 JAK-inhibitor – Upadacitinib
  - D11AH08 JAK-inhibitor – Abrocitinib).

Previous treatment for AD refers to treatment used within 6 months (i. e. 183 days) before treatment start with indication “treatment of atopic dermatitis related symptoms”.

The median for subgroup definitions will be calculated based on the respective analysis set used for subgroup analysis.

Frequencies of EASI 75 response (imputation according to B1_1 in Table 4-1) and summary statistics of percent change in EASI from baseline at Day 84 (Visit 7) (imputation according to C1_1 in Table 4-2) broken down by subgroups will be provided.

4.7 Interim Analyses

An interim analysis will *be performed using the data of all study participants having either regularly completed or prematurely discontinued study intervention at a certain cutoff date, when a minimum of 80% of patients are evaluable. The eligibility of the respective study participants for the analysis sets will be checked and documented before unblinding for the interim analysis. Study intervention of participants not having completed the intervention period by the time of the interim analysis will not be unblinded. Access to the treatment codes of the participants included in the interim analysis prior to final database release will be restricted to an unblinded team, whereas the rest of the team remains blinded. No individual data of the interim analysis will be disclosed to the blinded study team responsible for further study conduct. Details of the measures to restrict access to unblinding data will be described in detail in a dedicated study blinding plan. The further study conduct or study design will not be affected by the results of the interim analysis.*

An administrative interim analysis will be performed using the data of all study participants randomized until a certain date (02-AUG-2023). Approximately 90% of all participants have been randomized until this date and will be part of the interim analysis. The data cut-off for the interim analysis will be performed after all these participants have either regularly completed or prematurely discontinued study intervention and have primary efficacy data. The eligibility of the respective study participants for the analysis sets relevant for the interim analysis (SAS, FAS, PPS) will be checked and documented before unblinding for the interim analysis. Details of the measures to restrict access to unblinding data will be described in detail in a dedicated study blinding plan which will be signed before unblinding of the interim data.

A subset of the outputs planned for the final analysis will be provided for the interim analysis. The focus of the interim analysis will be the primary and secondary efficacy endpoints as well as the frequency of adverse events. The outputs will be provided in such a way that no unblinding on individual subject level will be possible. For summary statistics, only the number of subjects, the mean and standard deviation will be provided given that the number of participants per treatment group is at least 3. For frequency tables, number and percent of participants by treatment group will generally only be presented if there is at least one participant per treatment group with the event being counted or none in both treatment groups. In case of frequencies for related categories (e. g. several age categories) special care will be taken that there is no unblinding via total number of subjects per treatment group. For example, if there is one category with only one subject where frequencies by treatment group will not be presented the treatment of the subject could be derived by the numbers in the other categories by treatment group and total number of subjects by treatment group. In such a case the category with the next smallest total frequency will also only be presented with total frequency, not frequencies by treatment group. Total frequencies for both treatment groups combined can always be presented. Frequencies for the binary efficacy endpoints will be presented by treatment group to allow assessment of the efficacy of zabedosertib. However,

the individual post-dose data of these efficacy endpoints are stored in such a way that Bayer colleagues who will see the interim results either don't have access to this data or will not participate in further study conduct and blind data review meeting until unblinding for the final analysis.

4.8 Changes to Protocol-planned Analyses

According to the protocol a blinded monitoring of sample size assumptions was planned. This monitoring will not be applied because on the one hand the recruitment period turned out to be quite short (approximately 5 months) while on the other hand the study duration for a single participant is quite long (approximately 5 months). There is sufficient confidence in assumptions for sample size to go without further monitoring.

For the bayesian analysis of primary and binary secondary endpoints 80% and 95% credible intervals will be listed instead of 90 and 95% credible intervals to add one confidence interval corresponding to the clinical activity criterion.

5. Sample Size Determination

A Bayesian concept will be used for inference, which is based on the posterior probability that the research hypothesis is true. Therefore, it is the goal of this study to demonstrate that zabedosertib is superior to placebo with posterior probability > 90%, i.e. P($\pi_{BAY} > \pi_{Plac}$ | data) > 90% (='clinical activity' criterion).

The EASI 75 responder rate of 0.49 for assumed systemic standard of care (SoC) dupilumab was derived as the mean responder rate from a meta-analysis on the active treatment arms in two Phase 3 AD studies (Simpson et al, 2017), comprising 919 participants.

Non-informative Beta(1,1) priors will be used for the Bayesian inference.

For Placebo, a response rate of 16% was assumed. The statistical power was determined with SAS 9.4.

Assuming that the EASI 75 responder rate in zabedosertib-treated participants is 0.49, evaluable data from 57 participants (38 on zabedosertib, 19 on placebo) will provide at least 90% power to establish superiority of zabedosertib. Approximately 72 participants will be randomly assigned to the study to obtain 57 evaluable participants overall (2:1 randomization zabedosertib versus placebo).

6. Supporting Documentation

The number of participants enrolled, randomized and valid for the SAF, PPS, FAS, PKS, BAS will be summarized for each study site and overall. The number of participants randomized, and the number of sites will further be summarized by country. The number of participants with important protocol deviations and the number of screen failures will be summarized by country and study site.

In addition, an overview of all participants in the different analysis sets will be provided together with reasons for restrictions of validity (based on all participants randomized).

Furthermore, the important protocol deviations will be listed and summarized in a frequency table (based on all participants randomized).

An overview will be given of all participants who were enrolled, who were randomized, who were treated and who completed the study (based on all participants enrolled, i. e. signed the informed consent form). Completion status of the study will be presented for all epochs: screening (all participants enrolled), treatment, follow-up (based on all participants randomized), including reasons for not completing the respective epoch (incl. COVID-19-related discontinuations).

Participants affected by COVID-19 pandemic-related study disruption will be listed.

6.1 Demography and other baseline characteristics

Demographic data and baseline characteristics will be presented for the SAF, FAS, PPS.

Summary statistics will be provided for quantitative demographic data and baseline characteristics including:

- age,
- weight at baseline,
- height,
- body mass index (BMI) at baseline,
- body mass index (BMI) at baseline categories (<25, 25 - 30, >30)
- EASI at baseline,
- BSA affected by AD at baseline,
- weekly average of Peak Pruritus 0-10 NRS at baseline,
- disease duration at baseline,
- age of onset,
- number of AD flares requiring treatment in the last year before informed consent,
- number of skin infections requiring treatment in the last year before informed consent),
- IgE level at baseline.

Frequency tables will be provided for qualitative data including:

- age ≥18 – <40, ≥40 - ≤ 65,
- sex,
- race,
- ethnicity,
- vIGA=3 at baseline,
- vIGA=4 at baseline,
- age of onset categories (<18 / >=18 years)

- IgE level at baseline (<300 kU/L, 300-1000 kU/L, > 1000 kU/L),
- history of asthma, allergic rhinitis, allergic conjunctivitis (overall and by condition),
- frequency of AD flares requiring treatment in the last year before informed consent,
- frequency of skin infections requiring treatment in the last year before informed consent,
- country.

Results will be presented for each treatment group and in total. These summary statistics will also be provided for PKS and BAS.

6.2 Medical history

Medical occurrences that begin before the start of study intervention but after obtaining informed consent will be recorded on the Medical History/Current Medical Conditions section of the case report form (CRF) not the AE section (unless procedure related, these will be recorded as pre-treatment AEs, see Section 4.5.2).

Medical history findings will be summarized differentiating between prior and ongoing disorders at start of study intervention using medical dictionary for regulatory activities (MedDRA) terms using the most recent effective MedDRA version.

6.3 Prior and concomitant medication and procedures

For medications, the following definitions apply:

- Prior medication: Medication taken before start of the study intervention intake, (regardless of when it ended). According to the protocol only medication taken within 6 months (i. e. 183 days) before treatment start will be included.
- Concomitant medication: Medication taken during treatment period, i.e., between first and last study intervention intake (regardless of when it started or ended).
- Posterior medication: Medication started after end of treatment.

Prior, concomitant and posterior medication will be coded to Anatomical Therapeutic Chemical (ATC) classification codes according to the World Health Organization Drug Dictionary (WHO-DD, current version at the time of analysis). The number of participants taking prior, concomitant and posterior medication will be presented using ATC classes and subclasses. For posterior medication only participants will be considered who completed the follow-up period.

Furthermore, separate frequency tables will be provided for prior systemic and topical AD medication that ended prior to first study drug intake as well as concomitant and posterior systemic and topical AD medication as follows (respective ATC codes are given in brackets):

Systemic AD medication (route of administration: oral, intramuscular, subcutaneous, intravenous)

- Glucocorticoids for systemic use (H02AB, H02BX01)
- Calcineurin inhibitors (L04AD),
- Selective immunosuppressants (as listed below)
  - Mycophenolic acid (L04AA06),
  - Baricitinib (JAK inhibitor) (L04AA37),
  - Upadacitinib (JAK inhibitor) (L04AA44),
  - Abrocitinib (JAK inhibitor) (D11AH08)
- Other immunosuppressants (as listed below)
  - Methotrexate (L04AX03, L01BA01),
  - Azathioprine (purine antagonist) (L04AX01),
- Agents for dermatitis, excluding corticosteroids (as listed below) (D11AH, only systemic as listed below)
  - Dupilumab (IL-4 and IL-13 antibody) (D11AH05),
  - Tralokinumab (IL-13 antibody) (D11AH07),

Topical AD medication (route of administration: topical, cutaneous)

Corticosteroids for topical use (plain or combination with antiseptics/antibiotics/other)

- Corticosteroids, weak (group I) (D07AA, D07BA, D07CA, D07XA)
- Corticosteroids, moderate (group II) (D07AB, D07BB, D07CB, D07XB),
- Corticosteroids, potent (group III) (D07AC, D07BC, D07CC, D07XC),
- Corticosteroids, very potent (group IV) (D07AD, D07BD, D07CD, D07XD)

- Agents for dermatitis, excluding corticosteroids (as listed below) (D11AH, only topical as listed below)
  - Calcineurin inhibitors (D11AH01, D11AH02),
  - Cromoglicic acid (D11AH03),
  - Crisaborole (PDE4 inhibitor) (D11AH06),
  - Ruxolitinib (JAK inhibitor) (D11AH09)

In case of multiple possible potency groups for corticosteroids the highest possible potency will be considered.

Medication will be considered as AD medication if the indication for use is

- “rescue treatment: protocol specified”

or if the indication is

- “treatment of atopic dermatitis related symptoms”

or if the indication is

- “adverse event” and the adverse event is consistent with the indication AD
  Adverse events will be classified as consistent with AD or not during blind data

review by persons without access to any individual post-dose efficacy data. The following AEs will always be considered to be consistant with AD:
- PT: Dermatitis atopic,
- PT: Neurodermatitis,
- PT: Eczema.

Frequencies will be presented for each category and all subcategories.

Furthermore, a frequency table will be presented for relevant prior, concomitant and posterior procedures (other than those scheduled in the protocol). For posterior procedures only participants will be considered who completed the follow-up period.

For procedures the following definitions apply:

- Prior procedure: Procedure applied before start of the study intervention intake (only prior procedures within 6 months (i.e. 183 days) before treatment start will be considered),
- Concomitant procedure: Procedure applied during treatment period, i. e. between first and last study intervention intake,
- Posterior procedure: Procedure applied after end of treatment.

7. References

Biomedical Informatics and Regulatory Review Science (BIRRS) Team, Center for Drug Evaluation and Research (CDER). Standard safety tables and figures: integrated guide. Version date: August 2022.

Hanifin J, Thurston M, Omoto M, et al. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. Exp Dermatol. 2001 Feb;10(1):11–18.

Phan NQ, Blome C, Fritz F, et al. Assessment of pruritus intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronic pruritus. Acta Derm Venereol. 2012;92(5):502-507.

Simpson EL, Akinlade B, Ardeleanu M. Two Phase 3 Trials of Dupilumab versus Placebo in Atopic Dermatitis. N Engl J Med. 2017;376(11):1090-1091.

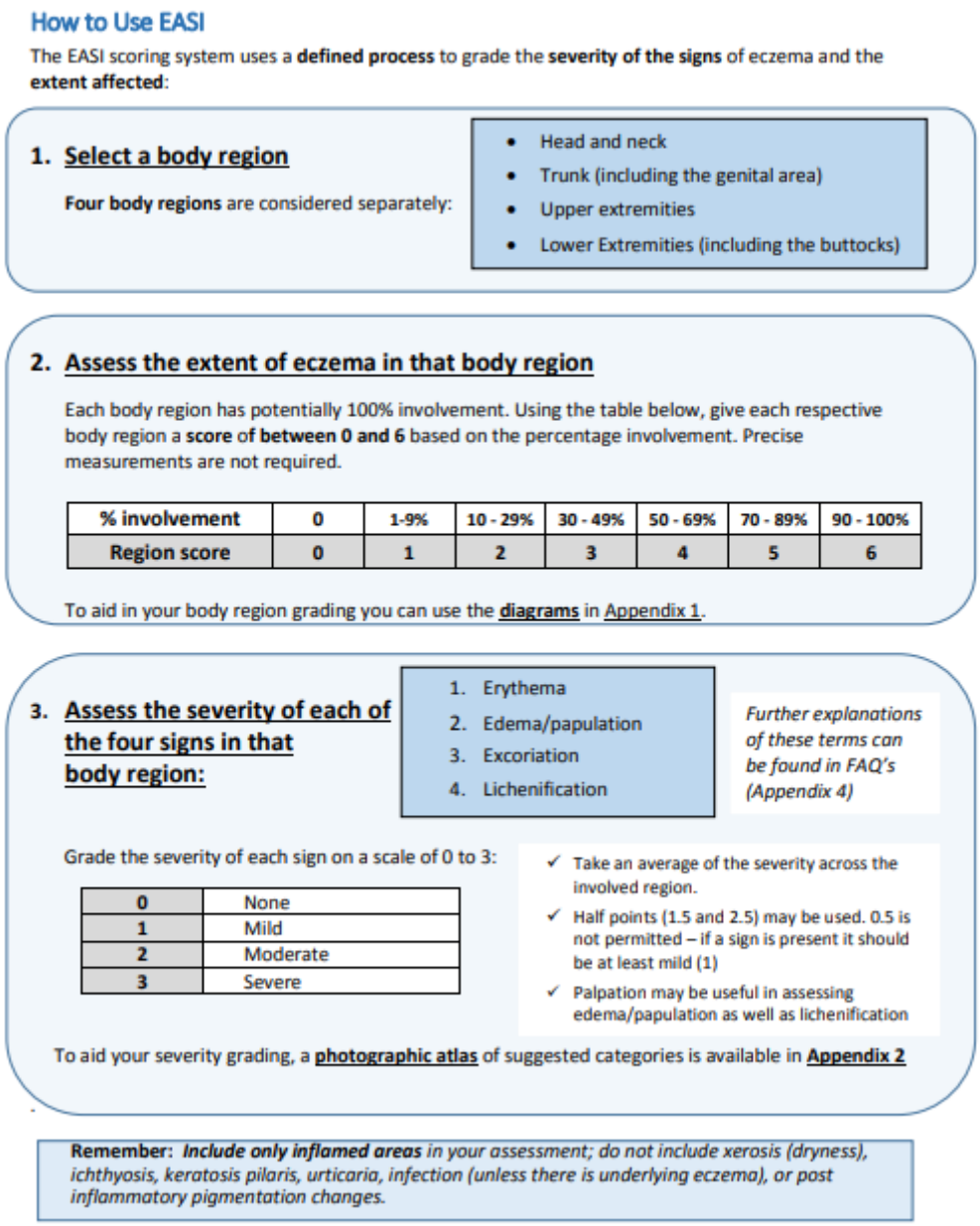
Zhou X, Reiter JP. A note on Bayesian inference after multiple imputation. The American Statistician. 2010; 64(2): 159-163.

8. Appendix

8.1 ClinROs scoring

8.1.1 Eczema Area and Severity Index (EASI)

Figure 1: EASI questionnaire instructions



How to Use EASI

The EASI scoring system uses a **defined process** to grade the **severity of the signs** of eczema and the **extent affected**:

1. Select a body region

Four body regions are considered separately:

- Head and neck
- Trunk (including the genital area)
- Upper extremities
- Lower Extremities (including the buttocks)

2. Assess the extent of eczema in that body region

Each body region has potentially 100% involvement. Using the table below, give each respective body region a **score of between 0 and 6** based on the percentage involvement. Precise measurements are not required.

% involvement	0	1-9%	10 - 29%	30 - 49%	50 - 69%	70 - 89%	90 - 100%
Region score	0	1	2	3	4	5	6

To aid in your body region grading you can use the **diagrams** in Appendix 1.

3. Assess the severity of each of the four signs in that body region:

1. Erythema
2. Edema/papulation
3. Excoriation
4. Lichenification

Further explanations of these terms can be found in FAQ's (Appendix 4)

Grade the severity of each sign on a scale of 0 to 3:

0	None
1	Mild
2	Moderate
3	Severe

- ✓ Take an average of the severity across the involved region.
- ✓ Half points (1.5 and 2.5) may be used. 0.5 is not permitted – if a sign is present it should be at least mild (1)
- ✓ Palpation may be useful in assessing edema/papulation as well as lichenification

To aid your severity grading, a **photographic atlas** of suggested categories is available in **Appendix 2**

Remember: ***Include only inflamed areas*** *in your assessment; do not include xerosis (dryness), ichthyosis, keratosis pilaris, urticaria, infection (unless there is underlying eczema), or post inflammatory pigmentation changes.*

easi-user-guide-dec-2016-v2.pdf (homeforeczema.org)

The assessed parameters are inserted into Table 8-1. The final EASI score ranges from 0-72.

Table 8-1: EASI scoring table

Body region	Erythema	Edema/ papulation	Excoriation	Lichenification	Area score	Multiplier	Score
Head/neck	(+	+	+	+)	x	x 0.1	
Trunk	(+	+	+	+)	x	x 0.3	
Upper extremities	(+	+	+	+)	x	x 0.2	
Lower extremities	(+	+	+	+)	x	x 0.4	
The final EASI score is the sum of the 4 region scores							——— (0-72)

8.1.2 Validated Investigator Global Assessment for Atopic Dermatitis (vIGA-AD)

The clinicians are asked to select that vIGA-AD score that best describes the overall appearance of the lesions at a given point in time. The scale's morphology descriptors are nonoverlapping, noncomparative categories with "clear" skin representing the absence of disease and focus given to having clearly distinct categories for ratings of "mild" versus "almost clear". A decrease in the vIGA-AD score relates to an improvement in signs and symptoms.

Instructions:

The IGA score is selected using the descriptors below that best describe the overall appearance of the lesions at a given time point. It is not necessary that all characteristics under Morphological Description be present.

Table 8-2: COA instrument: vIGA-AD

Score	Morphological Description
0 – Clear	No inflammatory signs of atopic dermatitis (no erythema, no induration/papulation, no lichenification, no oozing/crusting). Post-inflammatory hyperpigmentation and/or hypopigmentation may be present.
1 – Almost clear	Barely perceptible erythema, barely perceptible induration/papulation, and/or minimal lichenification. No oozing or crusting.
2 – Mild	Slight but definite erythema (pink), slight but definite induration/papulation, and/or slight but definite lichenification. No oozing or crusting.
3 – Moderate	Clearly perceptible erythema (dull red), clearly perceptible induration/papulation, and/or clearly perceptible lichenification. Oozing and crusting may be present.
4 – Severe	Marked erythema (deep or bright red), marked induration/papulation, and/or marked lichenification. Disease is widespread in extent. Oozing or crusting may be present.

Notes:
1. In indeterminate cases, please use extent to differentiate between scores.
For example:
• Patient with marked erythema (deep or bright red), marked papulation and/or marked lichenification that is limited in extent, will be considered “3 – Moderate”.
2. Excoriations should not be considered when assessing disease severity.

Validated-Investigator-Global-Assessment-Scale_vIGA-AD_2017.pdf (eczemacouncil.org)

8.1.3 Assessment of body surface area (BSA) affected by atopic dermatitis

Body surface area affected by AD will be assessed for each section of the body, using the “rule of nines”. The possible highest score for each region is:

Head and neck: 9%
Anterior trunk: 18%
Back: 18%
Upper limbs: 18%
Lower limbs: 36%
Genitals: 1%.

It will be reported as a percentage of all major body sections combined.

Before the start of the study at the individual site, all investigators will be trained on the correct assessment of BSA affected by AD.

8.2 PROs scales and questionnaires

Figure 2: Peak Pruritus 0-10 NRS

On a scale of 0 to 10, with 0 being 'no itch' and 10 being 'worst itch imaginable', how would you rate your itch at the worst moment during the previous 24 hours?

0	1	2	3	4	5	6	7	8	9	10
No itch										Worst itch imaginable

8.3 Assessment of intercurrent events

Identification of intercurrent event "use of rescue medication"

All medication on the concomitant medication page that is rescue medication according to the protocol and used at or after treatment start will be identified:

- topical corticosteroids (TCS)
  (Defined rescue medication for this study is the methylprednisolone aceponate (Advantan® cream). Only if Advantan cream is not available in the respective country, an alternative TCS with similar potency may be used (see WHO definition for potent corticosteroids group III, such as betamethasone, fluclorolone, desoximetasone, fluocinolone acetonide, budesonide, mometasone, beclomethasone, hydrocortisone aceponate))
- topical calcineurin inhibitors (TCI)
  (If TCSs are not tolerated or their use is not advisable, the topical TCIs pimecrolimus 1% (Elidel® cream) or tacrolimus (Protopic® 0.1% or 0.03% ointment) may be used.

Rescue medication will be identified by the following ATC codes if route of administration is topical or cutaneous (in case of multiple possible potency groups for corticosteroids the highest possible potency will be considered):

- Corticosteroids, potent (group III) (D07AC, D07BC, D07CC, D07XC),
- Corticosteroids, very potent (group IV) (D07AD, D07BD, D07CD, D07XD),
- Agents for dermatitis, excluding corticosteroids:
  - Calcineurin-inhibitors – Tacrolimus (D11AH01),
  - Calcineurin-inhibitors – Pimecrolimus (D11AH02)

The occurrence of the intercurrent event "use of rescue medication" will be applicable if the indication for use of an above listed topical medication is

- "rescue treatment: protocol specified"

or if the indication is

- "treatment of atopic dermatitis related symptoms"

or if the indication is

- "adverse event" and the adverse event is consistent with the indication AD
  Adverse events will be classified as consistent with AD or not during blind data review by persons without access to any individual post-dose efficacy data. The following AEs will always be considered to be consistant with AD:
  - PT: Dermatitis atopic,
  - PT: Neurodermatitis,
  - PT: Eczema.

If rescue medication is given for indication not consistent with AD this might also affect AD related efficacy scores. Thus, efficacy values assessed after rescue medication use at or after Day 56 (Visit 6) (or Day x (Visit x) for supplementary analyses) will be set to missing and the missing values will be imputed as described for the intercurrent event "discontinuation of study intervention due to other reasons".

Identification of intercurrent event "use of systemic standard of care"

All medication on the concomitant medication page that is systemic standard of care according to the protocol and was used at or after treatment start will be identified:

- *systemic corticosteroids,*
- *non-steroidal systemic immunosuppressive drugs (cyclosporine, methotrexate, MMF, azathioprine, baricitinib etc.),*
- *dupilumab,*
- *tralokinumab,*
- *phototherapy.*

Systemic SoC will be identified by the following ATC codes if route of administration is oral, intramuscular, subcutaneous or intravenous:

- H02AB/H02BX Glucocorticoids for systemic use
- L04AD: Calcineurin-inhibitors
  - L04AD 01 Ciclosporin
  - L04AD 02 Tacrolimus
  - L04AD 03 Voclosporin
- L04AA: selective immunosuppressants
  - L04AA06 MMF (Mycophenolic acid, Mycophenolate sodium, Mycophenolate acid, Mycophenolate mofetil)
  - L04AA37 JAK-inhibitor – Baricitinib
  - L04AA44 JAK-inhibitor – Upadacitinib
- D11AH08 JAK-inhibitor – Abrocitinib
- L04AX: other immunosuppressants
  - L04AX01 Purine antagonist – Azathioprine
  - L04AX03, L01BA01 Antineoplastic folic acid analogues – Methotrexate

- D11AH05, IL-4 and IL-13 monoclonal antibody – Dupilumab
- D11AH07 IL-13 antibodies – Tralokinumab
- L04AC Interleukin inhibitors used as immunosuppressants
  - L04AC IL-13 antibodies – Lebrikizumab not approved
  - L04AC IL-31 antibodies – Nemolizumab not approved
- Phototherapy (Psoralens for systemic (D05BA) or topical (D05AD) use before/together with phototherapy (PUVA))

The occurrence of the intercurrent event "use of systemic standard of care" will be applicable if the indication for use for the therapies listed above is

- "Treatment of atopic dermatitis related symptoms"

or if indication is

- "Adverse event" and the adverse event is consistent with the indication AD (Adverse events will be classified as consistent with AD or not during blind data review by persons without access to any individual post-dose efficacy data. The following AEs will always be considered to be consistant with AD:
  - PT: Dermatitis atopic,
  - PT: Neurodermatitis,
  - PT: Eczema).

If systemic standard of care medication is given for indication not consistent with AD this will also affect AD related efficacy scores. Thus, efficacy values assessed after systemic SoC use will be set to missing and the missing values will be imputed as described for the intercurrent event "discontinuation of study intervention due to other reasons".

<u>Identification of intercurrent event "Discontinuation of study intervention due to lack of efficacy"</u>

All reasons for treatment discontinuation will be reviewed. The occurrence of the intercurrent event "Discontinuation of treatment due to lack of efficacy" will be applicable if the underlying reason for discontinuation of treatment is

- "Adverse event" that reflects lack of efficacy (e.g. disease flare, worsening of AD) (Adverse events leading to discontinuation of treatment will be classified as reflecting lack of efficacy or not during blind data review by persons without access to any individual post-dose efficacy data. The following AEs will always be considered to reflect lack of efficacy)
  - PT: Dermatitis atopic with
    - LLT : Atopic dermatitis flare
    - LLT : Atopic flare-up
    - LLT : Dermatitis atopic aggravated
  - PT: Rebound atopic dermatitis
  - PT: Neurodermatitis with
    - LLT : Neurodermatitis aggravated

- PT: Eczema with
  - LLT : Eczema aggravated
  - LLT : Eczema exacerbated
  - LLT: Eczema
- PT: Rebound eczema

or if the reason is

- "Lack of efficacy" based on the investigator's assessment and there is no respective adverse event reported.

or if the reason is

- "Physician decision: switching to other therapy"
  This should be selected if any prohibited pharmacologically active non-study medication or phototherapy for AD is started and the study medication needs to be discontinued for this reason.

8.4 Handling of efficacy data (ClinRO and PRO) and imputation strategy

8.4.1 Data handling and visit windows for analysis

After occurrence of an intercurrent event (such as use of rescue medication or systemic SoC) the analysis will be performed according to the estimands. Data collected after use of topical rescue medication or systemic SoC are likely to show improvement in efficacy assessments. Such favourable values should not be included for analysis (unless stated otherwise, i.e. for the analysis of the effect of study intervention and rescue medication combined).

Therefore, the following first data handling step will be performed for participants who use systemic SoC or who use topical rescue medication at or after Day 56 (Visit 6) (or Day x (Visit x) for supplementary analyses):
All efficacy assessments (ClinRO or PRO) collected after (i.e. on the next day or later) first start of systemic SoC or start of topical rescue medication at or after Day 56 (Visit 6) (or Day x (Visit x) for supplementary analyses) will be set to missing (including FU visit). In case of use of rescue medication at Day 56 (Visit 6) (or Day x (Visit x) for supplementary analyses) which already started before that day the assessments after start of rescue medication will be set to missing as well.
This refers to use of any medication from the systemic SoC list and topical rescue medication list in Section 8.3 no matter if the medication was adjudicated as intercurrent event "use of systemic SoC" / "use of rescue medication at or after Day 56 (Visit 6)" as described in Section 8.3 or not.

Then the following data handling steps will be performed for all participants:

Efficacy assessments (i. e. ClinRO and ePRO weekly average of Peak Pruritus 0-10 NRS) at planned visits will be considered valid for analysis if they fall in the following visit windows (for Peak Pruritus the assessments at the actual visit days will be checked):

Table 8-3: Visit windows for analysis

Day 1 (Visit 3)	Day 1*
Day 14 (Visit 4)	[Day 10, Day 18]
Day 28 (Visit 5)	[Day 21, Day 35]
Day 56 (Visit 6)	[Day 49, Day 63]
Day 84 (Visit 7)	[Day 76, Day 92]
FU (Visit 8)	no restriction

* If the efficacy assessment is done at Day 1 (Visit 3) but treatment is started later, the assessment will also be acceptable.

Likewise, valid values from end of treatment visits and unscheduled visits can be assigned to the respective planned visits if they fall into the visit window as listed above. If several assessments fall into the same visit window the one closer to the planned visit day will be taken unless rescue medication was taken before at least one of the assessments. If there is still an assessment(s) without prior rescue medication use the one closest to the planned visit day will be selected. Otherwise, the one with greater distance to prior rescue medication use will be selected. (If there should be two assessments with equal distance to the planned visit date and not affected by rescue medication use then the one assessed at the regular visit will be taken.) Assessments from end of treatment visits after early discontinuation of study drug will only be considered for assignment to a planned visit if the end of treatment visit was not later than three days after last study drug intake.

Especially, it will be checked according to the criteria specified above if data from an early discontinuation visit or unscheduled visit before the use of rescue medication at Day 56 (Visit 6) (or Day x (Visit x) for supplementary analyses) or use of systemic SoC is available which fulfills the rules defined above. Likewise, in case of early discontinuation due to lack of efficacy or due to other reasons it will be checked if data from an early discontinuation visit or unscheduled visit before the intercurrent event is available which fulfills the rules defined above. If so, these values will be assigned to the respective planned visit.

Imputation of efficacy data for visits after an intercurrent event will be described in Section 8.4.2 and 8.4.3.

Note that events of “use of topical rescue medication” and “use of systemic standard of care” which are not considered intercurrent events as described in Section 8.3 (i. e. if the medication was not used for treatment of AD) will be handled and imputed as described for the intercurrent event “discontinuation of study intervention due to other reasons” (refer to Section 8.3).

Furthermore, if a regular visit at Day 56 (Visit 6) (or Day x (Visit x) for supplementary analyses) (i.e. a visit in the respective visit window for analysis as stated above) is changed into an end of treatment visit and rescue medication is taken at that visit this will be handled similarly to “use of rescue medication at or after Day 56” (or Day x (Visit x) for supplementary analyses) for analysis.

8.4.2 CCI

CCI

[redacted]

[redacted]

CCI

CCI

[redacted]

CCI

CCI

[redacted]

8.5 Reference ranges for laboratory parameters

Table 8-4 Reference ranges for laboratory parameters which differ by gender or age

Parameter	Sex	>= Age (years)	< Age (years)	Lower Limit	Upper limit	Unit
Glucose fasting	male	18	61	74	106	mg/dL
Glucose fasting	male	61	67	82	115	mg/dL
Glucose fasting	female	18	61	74	106	mg/dL
Glucose fasting	female	61	67	82	115	mg/dL
Glucose non-fasting	male	18	67	74	139	mg/dL
Glucose non-fasting	female	18	67	74	139	mg/dL
Uric acid (Urate)	male	18	67	3.4	7.0	mg/dL
Uric acid (Urate)	female	18	67	2,4	5,7	mg/dL
Calcium	male	18	61	2.15	2.50	mmol/L
Calcium	male	61	67	2,20	2.55	mmol/L
Calcium	female	18	61	2.15	2.50	mmol/L
Calcium	female	61	67	2,20	2.55	mmol/L
Creatinine	male	18	67	0.67	1.17	mg/dL
Creatinine	female	18	67	0,51	0.95	mg/dL
Aspartate aminotransferase (AST/GOT)	male	18	67	0	50	U/L
Aspartate aminotransferase (AST/GOT)	female	18	67	0	35	U/L
Alanine aminotransferase (ALT/GPT)	male	18	67	0	50	U/L

Alanine aminotransferase (ALT/GPT)	female	18	67	0	35	U/L
Gamma glutamyl transferase (Gamma GT)	male	18	67	0	71	U/L
Gamma glutamyl transferase (Gamma GT)	female	18	67	0	42	U/L
Lactate dehydrogenase (LDH)	male	18	67	135	225	U/L
Lactate dehydrogenase (LDH)	female	18	67	135	214	U/L
Alkaline phosphatase (ALP)	male	18	67	40	129	U/L
Alkaline phosphatase (ALP)	female	18	67	35	104	U/L
Creatine kinase (CK)	male	18	67	0	308	U/L
Creatine kinase (CK)	female	18	67	0	192	U/L
Hematocrit	male	18	65	39.5	50.5	%
Hematocrit	male	65	67	37.0	49.0	%
Hematocrit	female	18	50	35.5	45.0	%
Hematocrit	female	50	65	35.5	45.5	%
Hematocrit	female	65	67	35.0	45.5	%
Hemoglobin	male	18	65	13.5	17.2	g/dL
Hemoglobin	male	65	67	12.5	17.2	g/dL
Hemoglobin	female	18	50	12.0	15.4	g/dL

Hemoglobin	female	50	65	12.0	15.6	g/dL
Hemoglobin	female	65	67	11.8	15.8	g/dL
Erythrocytes (RBC)	male	18	65	4.30	5.75	10^{12}/L
Erythrocytes (RBC)	male	65	67	4.00	5.65	10^{12}/L
Erythrocytes (RBC)	female	18	50	3.90	5.15	10^{12}/L
Erythrocytes (RBC)	female	50	65	3.90	5.20	10^{12}/L
Erythrocytes (RBC)	female	65	67	3.85	5.20	10^{12}/L