

Clinical Development

CAIN457/Secukinumab/Cosentyx®

Clinical Trial Protocol CAIN457F2342 / NCT02404350

**A Phase III, randomized, double-blind, placebo controlled multi-center study of subcutaneous secukinumab (150 mg and 300 mg) in prefilled syringe to demonstrate efficacy (including inhibition of structural damage), safety, and tolerability up to 2 years in subjects with active psoriatic arthritis (FUTURE 5)**

Statistical Analysis Plan (SAP) Amendment 4

Author: [REDACTED]

Document type: SAP Documentation

Document status: Final

Release date: 15-Aug-2017

Number of pages: 74

Property of Novartis  
For business use only  
May not be used, divulged, published or otherwise disclosed  
without the consent of Novartis

**Document History – Changes compared to previous final version of SAP**

Date	Reason for update	Outcome for update	Section and title impacted (Current)
08-Aug-2017	Clarifying the analyses for between AIN457 comparisons	Amendment 03	2.6.3, 2.7.3
	Aligning terminology with protocol by replacing “no disease progression” with “no structural progression”		2.6.1, 2.7.1, 4.2.8
	Correct description of the planned “de facto”-type analyses		2.6.2.2
			[REDACTED]
			[REDACTED]
			[REDACTED]
	Correct typographical errors Update references		Throughout document 5
			[REDACTED]

**Table of contents**

Table of contents .....	3
List of abbreviations .....	5
1 Introduction .....	8
1.1 Study design.....	8
1.2 Study objectives and endpoints .....	9
1.2.1 Primary objective .....	9
1.2.2 Secondary objectives.....	9
[REDACTED]	10
2 Statistical methods.....	11
2.1 Data analysis general information .....	11
2.2 Analysis sets and treatment groups.....	12
2.2.1 Subgroup of interest .....	12
2.2.2 Treatment groups .....	12
2.3 Patient disposition, demographics and other baseline characteristics .....	13
2.3.1 Patient disposition .....	13
2.3.2 Background and demographic characteristics.....	14
2.3.3 Medical history.....	14
2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance).....	15
2.4.1 Study treatment / compliance.....	15
2.4.2 Prior, concomitant and post therapies .....	15
2.5 Analysis of the primary objective.....	16
2.5.1 Primary endpoint.....	16
2.5.2 Statistical hypothesis, model, and method of analysis .....	16
2.5.3 Handling of missing values/censoring/discontinuations.....	17
2.5.4 Supportive analyses.....	17
2.6 Analysis of secondary objectives.....	18
2.6.1 Secondary endpoints .....	18
2.6.2 Statistical hypothesis, model, and method of analysis .....	18
2.6.3 Handling of missing values/censoring/discontinuations .....	25
[REDACTED]	26
[REDACTED]	26
[REDACTED]	27
[REDACTED]	29
2.8 Safety analyses.....	30

2.8.1	Adverse events (AEs).....	30
2.8.2	Laboratory data .....	31
2.8.3	Vital signs.....	34
2.8.4	Electrocardiogram (ECG) .....	35
2.8.5	[REDACTED]	35
2.8.6	Compound specific safety evaluation .....	35
	[REDACTED]	35
	[REDACTED]	36
	[REDACTED]	37
	[REDACTED]	38
	[REDACTED]	39
2.13	Interim analysis.....	54
3	Sample size calculation .....	54
4	Appendix .....	55
4.1	Visit Windows and Cut-off Dates.....	55
4.1.1	Visit windows.....	55
4.1.2	Cut-off dates.....	60
4.2	Detailed on implementation of statistical methodology and assumptions.....	60
4.2.1	Analysis of continuous data .....	60
4.2.2	Analysis of binary (and categorical) data.....	61
4.2.3	Multiple Imputation .....	63
4.2.4	Crude incidence and relativ risk estimates.....	66
4.2.5	Exposure adjusted incidence rate and related risk estimates .....	66
4.2.6	[REDACTED]	67
4.2.7	Details for Total vdH-mTSS analysis .....	69
	[REDACTED]	71
5	Reference .....	72

## List of abbreviations

ACR	American College of Rheumatology
AE	Adverse event
ALT/SGPT	Alanine aminotransferase/serum glutamic pyruvic transaminase
Anti-CCP	Anti-cyclic citrullinated peptide
AST/SGOT	Aspartate aminotransferase/serum glutamic oxaloacetic transaminase
ATC	Anatomical Therapeutic Classification
AUC	Area Under the Curve
bid	bis in diem/twice a day
BMI	Body Mass Index
BP	Blood Pressure
BSL	Baseline
CASPAR	CIASification criteria for Psoriatic ARthritis
CRF	Case Report/Record Form (paper or electronic)
CRO	Contract Research Organization
CRP/hsCRP	C-reactive protein / high sensitivity C-reactive protein
CSR	Clinical Study Report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
DAS	Disease Activity Score
DLQI	Dermatology Life Quality Index
DMARD	Disease Modifying Antirheumatic Drug
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	Electronic Case Report/Record Form
ESR	Erythrocyte Sedimentation Rate
EULAR	European League Against Rheumatism
FAS	Full Analysis Set
FDA	Food and Drug Administration
FSH	Follicle Stimulating Hormone
GCP	Good Clinical Practice
GEE	Generalized Estimating Equation
HAQ-DI	Health Assessment Questionnaire – Disability Index
HDL	High Density Lipoprotein
HRQoL	Health-related Quality of Life
hsCRP	High sensitivity C-Reaction Protein
IL	Interleukin
i.v.	Intravenous(ly)
IVR	Interactive Voice Response

IWR	Interactive Web Response
LDL	Low Density Lipoprotein
LLN	Lower Limit Normal
LLOQ	Lower Limit of Quantification
MAR	Missing at Random
MedDRA	Medical Dictionary for Drug Regulatory Affairs
mmHg	Millimeter mercury
MMRM	Mixed-effects Repeated Measures Model
MRI	Magnetic Resonance Imaging
MTX	Methotrexate
NSAID	Non-steroidal andti-inflammatory drug
PASI	Psoriatic Area and Severity Index
PCS	Physical Component Summary
PoC	Proof of Concept
PRO	Patient-reported Outcomes
PsA	Psoriatic Arthritis
QoL	Quality of Life
RA	Rheumatoid Arthritis
RAP	Report and Analysis Process
RBC	Red Blood Cells
RF	Rheumatoid Factor
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
s.c.	Subcutaneous(ly)
SCE	Summary of Clinical Efficacy
SCS	Summary of Clinical Safety
SD	Standard Deviation
SJC	Swollen Joint Count
SNP	Single Nucleotide polymorphism
SpA	Spondyloarthritis
SOC	System Organ Class
TFLs	Tables, Figures, Listings

---

TJC	Tender Joint Count
TNF	Tumor Necrosis Factor
TNF-IR	TNF $\alpha$ Inhibitor Inadequate Responders
ULN	Upper Limit of Normal
VAS	Visual Analog Scale
vdH-mTSS	van der Heijde modified total Sharpe Score
WBC	White Blood Cells
WHO	World Health Organization

## **1 Introduction**

Data will be analyzed by Novartis according to the data analysis section 9 of the clinical study protocol. The statistical methodology is described below and any deviations from the protocol are documented. Additional detailed information regarding the analysis methodology is contained in the Appendix section.

### **1.1 Study design**

This multicenter study uses a randomized, double-blind, placebo-controlled, parallel-group design. A screening period (SCR) running up to 10 weeks before randomization will be used to assess subject eligibility followed by 104 weeks of treatment.

At BSL approximately 990 subjects whose eligibility is confirmed will be randomized to one of four treatment groups in 2:2:2:3 ratio:

- Group 1 - secukinumab 150 mg s.c. without loading regimen
- Group 2 - secukinumab 150 mg s.c. with loading dose regimen
- Group 3 - secukinumab 300 mg s.c. with loading dose regimen
- Group 4 - Placebo s.c.

Subjects in Group 4 (Placebo) will be randomized (1:1) at BSL to two different treatment sequences:

Placebo s.c. till Week 16/24 followed by secukinumab 150 mg s.c. every 4 weeks starting at Week 16/24.

Placebo s.c. till Week 16/24 followed by secukinumab 300 mg s.c. every 4 weeks starting at Week 16/24.

**At randomization**, subjects will be stratified on the basis of previous anti-TNF therapy as TNF $\alpha$  inhibitor naïve (TNF-naïve) or TNF $\alpha$  inhibitor inadequate responders (TNF-IR).

At each study treatment visit, one (for secukinumab 150 mg) or two (for secukinumab 300 mg) s.c. injections in the form of PFS will be administered, since secukinumab is available in 1.0 mL (150 mg) PFSs. Placebo to secukinumab is also available in 1.0 mL to match the active drug.

**At Week 16**, subjects who have been randomized to secukinumab groups at BSL (Groups 1-3) will be classified as either responders ( $\geq 20\%$  improvement from BSL in both tender joint count (TJC) and swollen joint counts (SJC)) or non-responders ( $< 20\%$  improvement from BSL TJC or SJC), however, they will continue on the same treatment irrespective of their response status.

**At Week 16**, subjects who have been randomized to placebo at BSL (Group 4) will be classified as either responders ( $\geq 20\%$  improvement from BSL in both TJC and SJC) or non-responders ( $< 20\%$  improvement from BSL TJC or SJC):

- Subjects who are **non-responders** will receive either secukinumab 150 mg or 300 mg s.c. every 4 weeks starting at Week 16 (as dictated by treatment sequence assigned to these subjects at BSL).

- Subjects who are **responders** will continue to receive placebo every 4 weeks. At Week 24, these subjects will receive either secukinumab 150 mg s.c. or 300 mg s.c. every 4 weeks. (as dictated by treatment sequence assigned to these subjects at BSL).

**After Week 24**, all assessments related to the primary and key secondary objectives will have been performed.

**After the Week 52 database lock and analyses** have been completed, site personnel and subjects will be unblinded to the original randomized treatment (sequence) assignment at randomization. In addition, treatment will be given open-label in order to eliminate the placebo injection. The subject will continue to receive the same active dose of secukinumab as open-label treatment administered until Week 100.

## **1.2 Study objectives and endpoints**

### **1.2.1 Primary objective**

To demonstrate that the efficacy of secukinumab 150 mg s.c. (with or without loading regimen), or 300 mg s.c. with loading regimen, at Week 16 is superior to placebo based on proportion of subjects with active PsA achieving American College of Rheumatology 20 (ACR20) response.

### **1.2.2 Secondary objectives**

To evaluate:

The change from baseline at Week 24 with secukinumab 150 mg (with or without loading regimen), or 300 mg (with loading regimen) compared with placebo for **joint/bone structural damage (using van der Heijde modified total Sharp score (mTSS))**.

The efficacy of secukinumab 150 mg (with or without loading regimen), or 300 mg (with loading regimen) at Week 16 compared with placebo based on the proportion of subjects achieving **Psoriatic Area and Severity Index 75 (PASI75)** response.

The efficacy of secukinumab 150 mg (with or without loading regimen), or 300 mg (with loading regimen) at Week 16 compared with placebo based on the proportion of subjects achieving **Psoriatic Area and Severity Index 90 (PASI90)** response.

The efficacy of secukinumab 150 mg (with or without loading regimen), or 300 mg (with loading regimen), at Week 16 compared with placebo based on the proportion of subjects achieving an **ACR50** response.

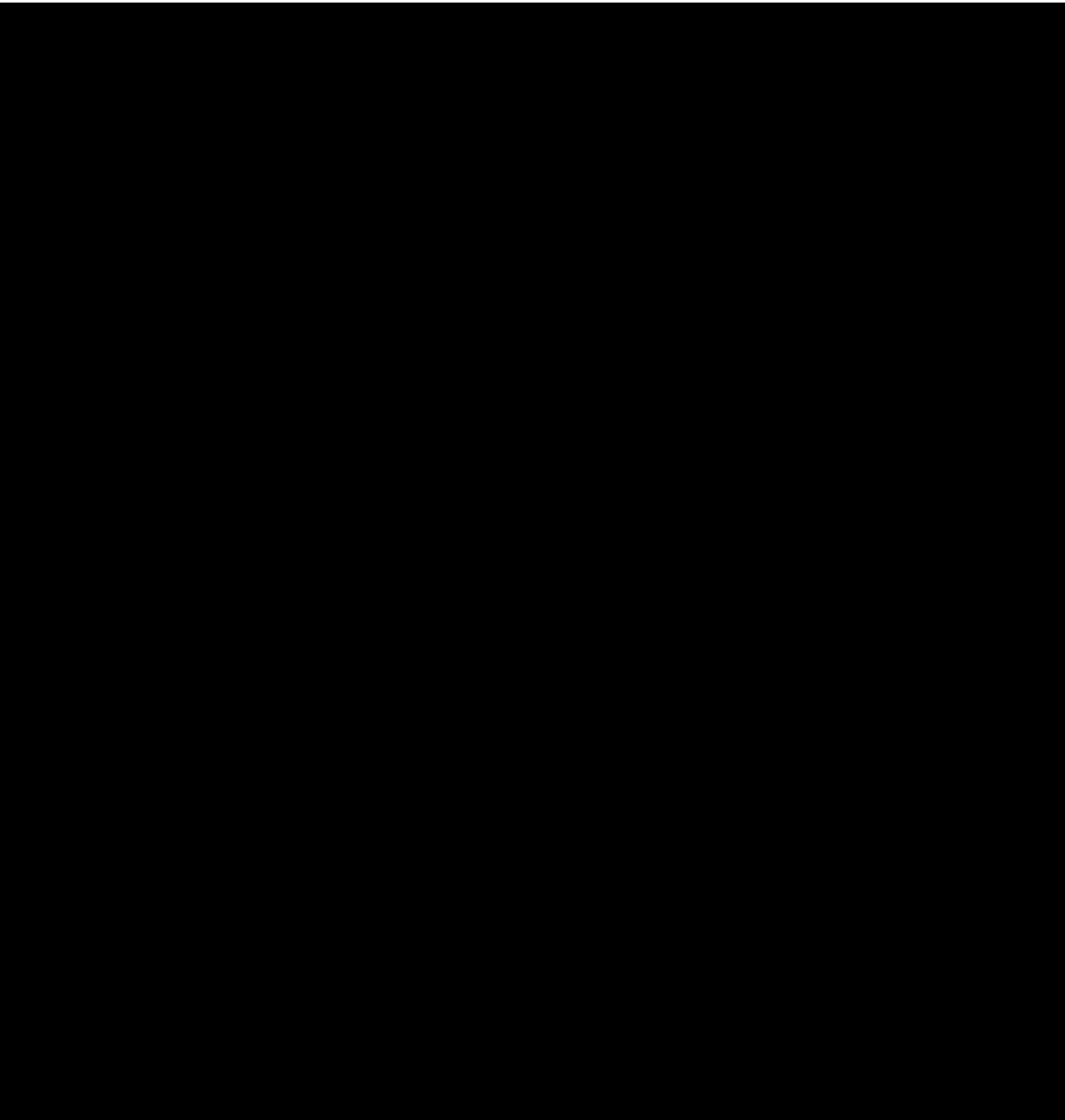
The improvement on secukinumab 150 mg (with or without loading regimen), or 300 mg (with loading regimen), at Week 16 compared with placebo for the disease activity assessed by the changes in **HAQ-DI<sup>®</sup>** relative to baseline.

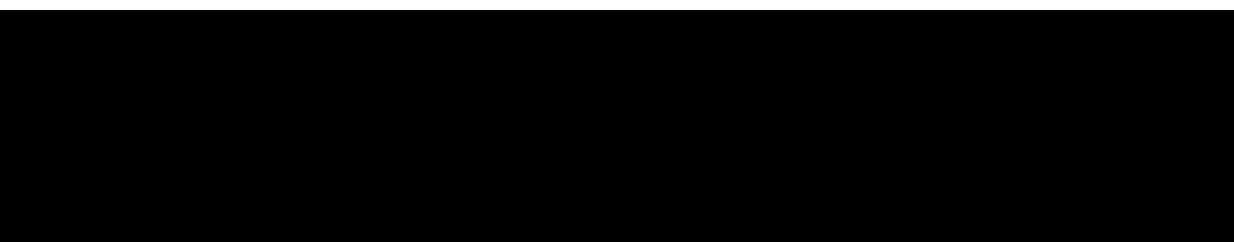
The improvement on secukinumab 150 mg (with or without loading regimen), or 300 mg (with loading regimen) at Week 16 compared with placebo for the disease activity assessed by the changes in **Disease Activity Score for 28 joints (DAS28-CRP)** (utilizing hsCRP) relative to baseline.

The efficacy of secukinumab pooled regimen (150 mg with or without loading regimen, and 300 mg with loading regimen) at Week 16 compared with placebo based on the proportion of subjects with **enthesitis** in the subset of subjects who have enthesitis at BSL.

The efficacy of secukinumab pooled regimen (150 mg with or without loading regimen, and 300 mg with loading regimen) at Week 16 compared with placebo based on the proportion of subjects with **dactylitis** in the subset of subjects who have dactylitis at BSL.

Overall safety and tolerability of secukinumab.





## **2 Statistical methods**

### **2.1 Data analysis general information**

Summary statistics for continuous variables include N, mean, standard deviation, minimum, lower quartile, median, upper quartile, and maximum. For binary or discrete variables the absolute number of subjects in each category and relative frequencies will be provided.

Unless otherwise specified, p-values will be presented as 2-sided p-values and the type I error rate (alpha) will be 5%.

Inferential efficacy comparisons with placebo will generally focus on the first 16-weeks of treatment except for radiographic assessment which will focus on the first 24-weeks of treatment unless otherwise specified.

Efficacy and safety data for the placebo-controlled period (or the entire treatment period as appropriate) will be presented by randomized treatment groups. Subjects may be included in more than one treatment group for some analyses (e.g. exposure-adjusted AEs over the entire treatment period).

Note that the treatment groups for a subject may differ depending on the time period of the analysis and whether one assesses the subject for efficacy or safety (see [Section 2.2.2](#) for details).

Data may also be presented by a combination of the 'original' and 'switch' treatment groups. These treatment groups represent the treatment combinations the subjects experience over the course of the entire trial.

Subjects will receive treatment at BSL, Weeks 1, 2 and 3, followed by treatment every 4 weeks starting with Week 4 through Week 104 (with last dose at Week 100).

### **Comparative efficacy data**

Comparative efficacy analyses (i.e. inferential efficacy comparisons with placebo) will focus on the time period when both active drug and the placebo are given in a manner suitable for making comparisons (e.g. double-blind). For AIN457F2342 this is the first 16/24-weeks of treatment depending on the endpoint considered. Comparative efficacy will be performed based on the FAS population using the randomized treatment. After week 24, the active secukinumab regimens will be compared using confidence intervals on the FAS population using treatment sequence.

## **Efficacy data following treatment switch**

Data will also be presented after Week 24, by a combination of the ‘original’ and ‘switch’ treatment groups and will be referred to as treatment sequence. These treatment sequences represent the treatment combinations the subjects experience over the course of the entire trial in case of treatment switch (e.g., placebo patients who are reassigned to 150 or 300 mg secukinumab at Week 16 (rescued) or Week 24).

All listings will be presented by treatment sequence.

## **2.2 Analysis sets and treatment groups**

The following analysis sets will be used for the data analysis.

**Randomized set:** The randomized set will be defined as all subjects who were randomized. Unless otherwise specified, mis-randomized subjects (mis-randomized in Interactive Voice Response (IVR)) will be excluded from the randomized set.

Mis-randomized subjects are defined as those subjects who were mistakenly randomized into the IVR prior to the site confirming all eligibility criteria had been met and to whom no study medication was given. Mis-randomized patients are treated as screen failures.

**Full analysis set (FAS):** The FAS will be comprised of all subjects from the randomized set to whom study treatment has been assigned. Following the intent-to-treat principle, subjects will be analyzed according to the treatment assigned to at randomization, but actual anti-TNF status.

**Dactylitis subset:** The dactylitis subset will include all FAS subjects who have dactylitis at baseline.

**Enthesitis subset:** The enthesitis subset will include all FAS subjects who have enthesitis at baseline.

**Psoriasis subset:** The psoriasis subset will include all FAS subject who have  $\geq 3\%$  of the body surface area (BSA) affected by psoriatic skin involvement at baseline.

**Nail subset:** The nail subset will include all FAS subject who have psoriasis currently in nails at baseline.

**Up-titration subset:** Up-titration subset will include all FAS subjects who have been up-titrated to secukinumab 300mg in the second year of the study.

**Safety set:** The safety set includes all subjects who took at least one dose of study treatment during the treatment period. Subjects will be evaluated according to treatment received.

### **2.2.1 Subgroup of interest**

The primary endpoint(s) and secondary endpoints will be evaluated for TNF-alpha inhibitor status.

### **2.2.2 Treatment groups**

The summaries by treatment will be performed by the randomized treatment or treatment sequence. For some safety summaries (e.g. exposure-adjusted) the ‘switch’ treatment may be summarized separately.

- Randomized treatment:
  - AIN457 150 mg No Load
  - AIN457 150 mg
  - AIN457 300 mg
  - Placebo
- Treatment sequence:
  - AIN457 150 mg No Load
    - Stay with 150 mg
    - Up-titrated to 300 mg
  - AIN457 150 mg
    - Stay with 150 mg
    - Up-titrated to 300 mg
  - AIN457 300 mg
  - Placebo Non-responder – AIN457 150 mg
    - Stay with 150 mg
    - Up-titrated to 300 mg
  - Placebo Non-responder – AIN457 300 mg
  - Placebo Responder – AIN457 150 mg
    - Stay with 150 mg
    - Up-titrated to 300 mg
  - Placebo Responder – AIN457 300 mg

*Note:* up-titration subgroup analysis will be performed for the final Week 112 DBL only.

Up-titrated subgroup will include those subjects who are up-titrated to AIN457 300 mg in the second year. Additional sequences could be reported such as Any AIN457; Any AIN457 150 mg; Any AIN457 300 mg if applicable.

## **2.3 Patient disposition, demographics and other baseline characteristics**

### **2.3.1 Patient disposition**

The number of subjects screened will be presented. In addition, the reasons for screen failures will be provided. The number and percentage of subjects in the randomized set who completed the study periods and who discontinued the study prematurely (including the reason for discontinuation) will be presented at the end of each treatment period (e.g. Week 24, 52, 104), if appropriate, for each treatment group and all subjects.

The number and percentage of patients who meet the rescue (treatment switch) criteria at week 16 will be presented.

For each protocol deviation (PD), the number and percentage of subjects for whom the PD applies will be tabulated.

### **2.3.2 Background and demographic characteristics**

The following common background and demographic variables, if collected, will be analyzed:

#### **Continuous variables:**

- Age (which is derived from date of birth and the screening assessment date)
- Height
- Weight
- Body mass index (BMI) = (body weight in kilograms) / (height in meters)<sup>2</sup>

#### **Categorical variables:**

- Age categories (<65 years, 65 years and older, 75 years and older)
- Gender
- Race
- Ethnicity
- Smoking status at baseline

The following disease specific baseline characteristics and history of disease will be summarized and also by concomitant TNF treatment:

- CASPAR total score, TNF $\alpha$  history (naive or inadequate responder, number of prior therapies), MTX use (yes or no) and dose of MTX or other DMARD at randomization, time since first diagnosis of PsA (years), ACR components, DAS28-CRP, DAS28-ESR, presence of enthesitis, presence of dactylitis, proportion of patients with: psoriasis of hands and feet, psoriasis of the nail, psoriasis  $\geq 3\%$  of BSA and number of prior biologic PsA therapies.

Unless otherwise specified, summary statistics will be presented for continuous variables for each treatment group and for all subjects (total) in the randomized set. The number and percentage of subjects in each category will be presented for categorical variables for each treatment group and all subjects (total) in the randomized set.

### **2.3.3 Medical history**

Any condition entered on the Relevant medical history / current medical conditions CRF will be coded using the MedDRA dictionary. They will be summarized by system organ class (SOC) and preferred term (PT) of the MedDRA dictionary. Summaries for cardiovascular medical history and psoriasis history will be provided as well. The number of cardiovascular risk factors that each subject has will also be summarized by treatment group. If it is unknown whether or not a subject currently or previously experienced a specific cardiovascular risk factor, it will be assumed that cardiovascular risk factor is not present for that subject

Smoking history will be summarized by treatment group.

Unless otherwise specified, analyses will be based on the randomized set.

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

### **2.4.1 Study treatment / compliance**

The analysis of study treatment data will be based on the safety set. The duration of exposure to study treatment will also be summarized by treatment group. In addition, the number and percentage of subjects with cumulative exposure levels (e.g. any exposure,  $\geq 1$  week,  $\geq 2$  weeks,  $\geq 3$  weeks,  $\geq 4$  weeks,  $\geq 8$  weeks, etc.) will be presented.

Duration of exposure of a treatment will be defined as the time from first dose of the treatment to the time of treatment switch (for subjects who switch treatment) or minimum of (last dose of the treatment + 84 days) and (last visit date). Patients who switch treatment during the study (e.g. from placebo to active treatment or between active treatment groups) will have exposure to both medications/doses using the appropriate start and stop dates.

Duration of exposure (years) = duration of exposure (days) / 365.25

Duration of exposure (100 subject years) = duration of exposure (years) / 100

The analyses of duration of exposure described above will be done for the entire study treatment period.

### **2.4.2 Prior, concomitant and post therapies**

Prior and concomitant medications will be summarized in separate tables by treatment group.

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Any medication given at least once between the day of first dose of study treatment and the date of within 84 days after last dose will be a concomitant medication, including those which were started pre-baseline and continued into the period where study treatment is administered.

Medications will be presented in alphabetical order, by Anatomical Therapeutic Classification (ATC) codes and grouped by anatomical main group. Tables will show the overall number and percentage of subjects receiving at least one treatment of a particular ATC code and at least one treatment in a particular anatomical main group.

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

The number and percentage of subjects receiving prior and concomitant psoriatic arthritis therapy will be presented by randomized treatment group as well as the reasons for stopping their therapies (primary lack of efficacy, secondary lack of efficacy, lack of tolerability, other).

Prior or concomitant medication will be identified by comparing recorded or imputed start and end dates of medication taken to the reference start date.

Prior anti-TNF PsA therapy is defined as one of the following prior medication: ADALIMUMAB, ETANERCEPT, GOLIMUMAB, INFliximab, and CERTOLIZUMAB PEGOL.

The number and rates of prior anti-TNF PsA therapies ( $=0/=1/>=2$ ) will be presented by randomized treatment.

## **2.5 Analysis of the primary objective**

### **2.5.1 Primary endpoint**

The primary efficacy variable will be ACR20 response at Week 16. The analysis of the primary efficacy variable will be based on the FAS. Primarily, CRP will be used instead of ESR to calculate ACR response; ESR will only be used in the event CRP is missing.

### **2.5.2 Statistical hypothesis, model, and method of analysis**

The statistical hypothesis for ACR20 being tested is that there is no difference in the proportion of subjects fulfilling the ACR20 criteria at Week 16 in the secukinumab regimens versus placebo regimen.

Let  $p_j$  denote the proportion of ACR20 responders at Week 16 for treatment regimens  $j$ ,  $j=0, 1, 2, 3$  where

- 0 corresponds to placebo regimen,
- 1 corresponds to secukinumab 300 mg,
- 2 corresponds to secukinumab 150 mg,
- 3 corresponds to secukinumab 150 mg no load,

In statistical terms,  $H_0: p_j = p_0$ ,  $H_{A_j}: p_j \neq p_0$ , for the  $j^{\text{th}}$  secukinumab regimen, i.e.

$H_1:$  secukinumab 300 mg is not different to placebo regimen with respect to signs and symptoms (ACR20 response) at Week 16

$H_2:$  secukinumab 150 mg is not different to placebo regimen with respect to signs and symptoms (ACR20 response) at Week 16

$H_3:$  secukinumab 150 mg no load is not different to placebo regimen with respect to signs and symptoms (ACR20 response) at Week 16

The primary endpoint of ACR20 at Week 16 in the FAS will be evaluated using a logistic regression with treatment and randomization stratum (TNF $\alpha$  status -naive or IR) as factors and weight as a covariate. Odds ratios will be computed for comparisons of secukinumab regimens versus placebo regimen utilizing the logistic regression model fitted.

The primary estimand of ACR20 response is the odds-ratio (each secukinumab dose vs. placebo) at week 16 in the FAS population that would be observed if all patients dropping out or being unblinded before week 16 or having with missing ACR20 data at week 16 were non-responders.

The overall testing strategy is explained in [Section 2.6.2.1](#).

## **2.5.3 Handling of missing values/censoring/discontinuations**

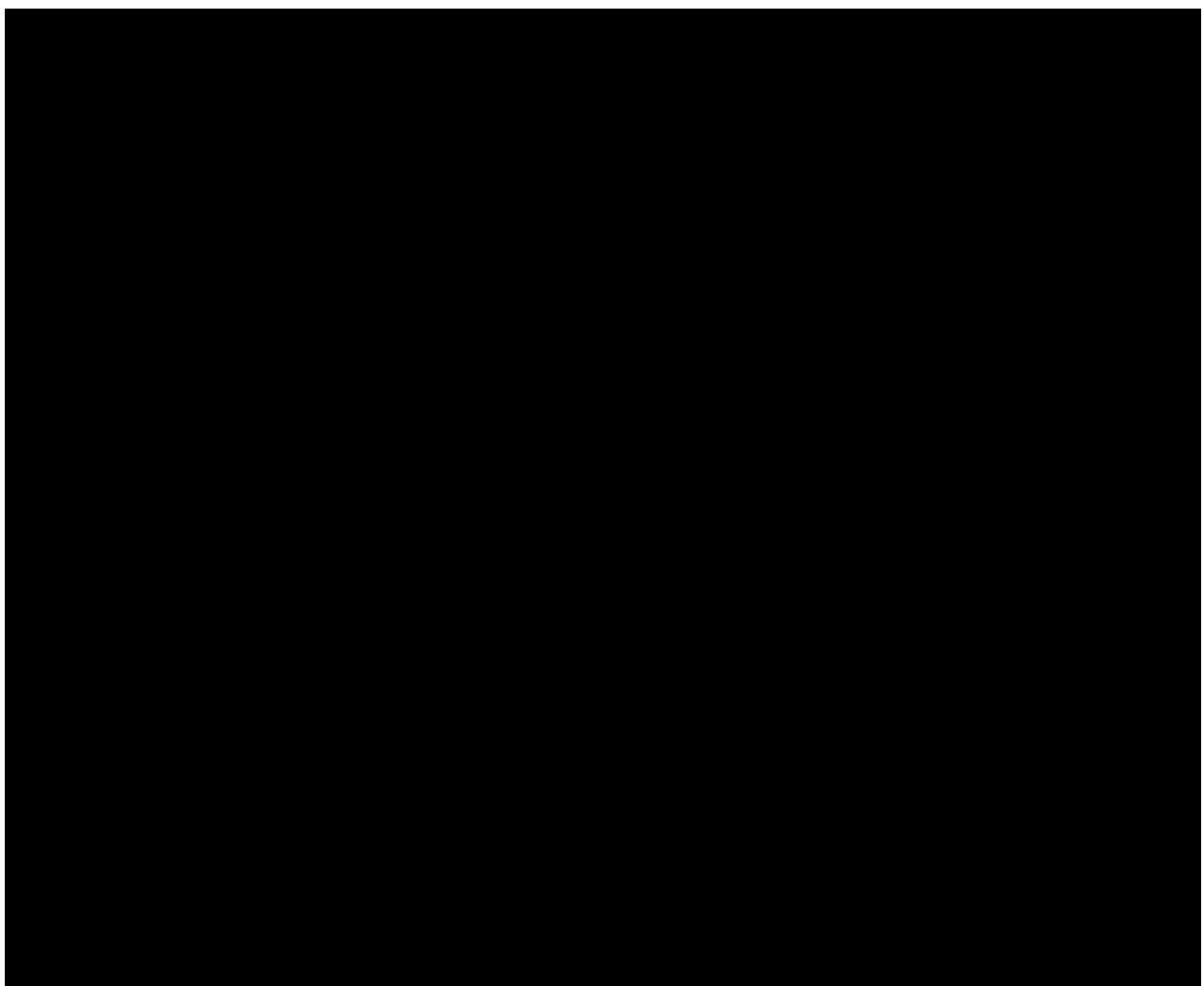
### **Missing data**

Missing data for ACR20 response for data up to Week 16 will be handled as follows:

- Subjects who drop out of the trial for any reason will be considered non-responders from the time they drop out through Week 16.
- Subjects who do not have the required data at baseline and at the specific time point to compute response will be classified as non-responders.

### **Premature unblinding**

Patients who were unblinded prior to the scheduled timepoint will be considered non-responders from the time of unblinding up to the end of the placebo-controlled period (Week 24). The primary analysis will use this non-responder imputation.



## **2.6 Analysis of secondary objectives**

### **2.6.1 Secondary endpoints**

The secondary efficacy variables are listed below. Secondary efficacy variables will be analyzed using the FAS population unless otherwise specified.

- Change from baseline in van der Heijde modified total Sharp score at Week 24
- PASI75 response at Week 16
- PASI90 response at Week 16
- ACR50 response at Week 16
- Change from baseline in HAQ-DI<sup>©</sup> score at Week 16
- Change from baseline in DAS28-CRP at Week 16
- Presence of enthesitis at Week 16
- Presence of dactylitis at Week 16

### **2.6.2 Statistical hypothesis, model, and method of analysis**

#### **2.6.2.1 Overall Testing strategy to control type I error**

The following hypotheses will be included in the testing strategy, and type-I-errors will be set such that a family-wise type-I-error of 5% is kept:

##### **Primary objectives:**

- H<sub>1</sub>: secukinumab 300 mg is not different to placebo regimen with respect to ACR20 response at Week 16
- H<sub>2</sub>: secukinumab 150 mg is not different to placebo regimen with respect to ACR20 response at Week 16
- H<sub>3</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to ACR20 response at Week 16

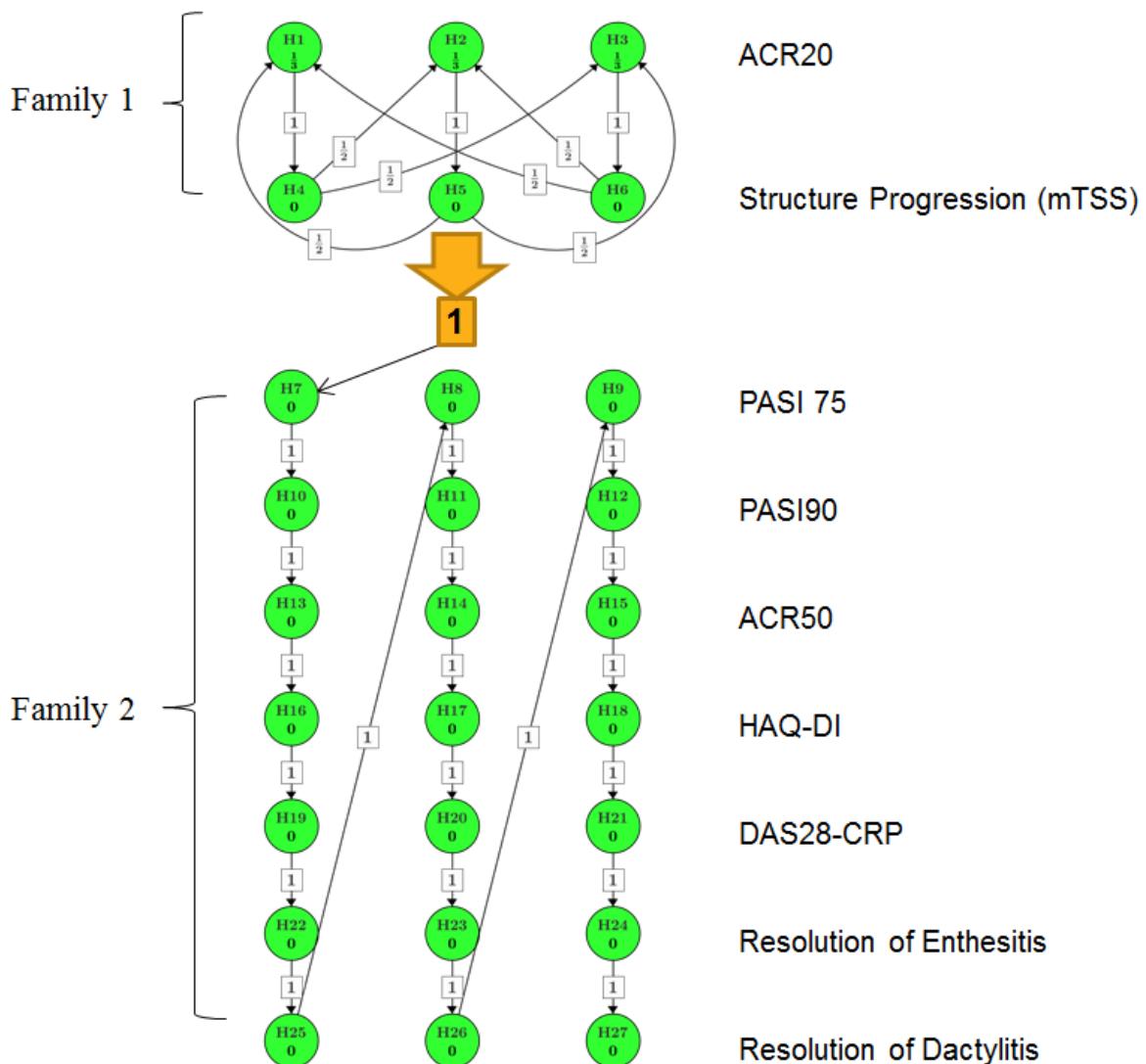
##### **Secondary objectives:**

- H<sub>4</sub>: secukinumab 300 mg is not different to placebo regimen with respect to structural damage (van der Heijde modified total Sharp score) at Week 24
- H<sub>5</sub>: secukinumab 150 mg is not different to placebo regimen with respect to structural damage (van der Heijde modified total Sharp score) at Week 24
- H<sub>6</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to structural damage (van der Heijde modified total Sharp score) at Week 24
- H<sub>7</sub>: secukinumab 300 mg is not different to placebo regimen with respect to PASI75 response at Week 16 in the subset of subjects who have  $\geq 3\%$  skin involvement with psoriasis
- H<sub>8</sub>: secukinumab 150 mg is not different to placebo regimen with respect to PASI75 response at Week 16 in the subset of subjects who have  $\geq 3\%$  skin involvement with psoriasis

- H<sub>9</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to PASI75 response at Week 16 in the subset of subjects who have  $\geq 3\%$  skin involvement with psoriasis
- H<sub>10</sub>: secukinumab 300 mg is not different to placebo regimen with respect to PASI90 response at Week 16 in the subset of subjects who have  $\geq 3\%$  skin involvement with psoriasis
- H<sub>11</sub>: secukinumab 150 mg is not different to placebo regimen with respect to PASI90 response at Week 16 in the subset of subjects who have  $\geq 3\%$  skin involvement with psoriasis
- H<sub>12</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to PASI90 response at Week 16 in the subset of subjects who have  $\geq 3\%$  skin involvement with psoriasis
- H<sub>13</sub>: secukinumab 300 mg is not different to placebo regimen with respect to ACR50 response at Week 16
- H<sub>14</sub>: secukinumab 150 mg is not different to placebo regimen with respect to ACR50 response at Week 16
- H<sub>15</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to ACR50 response at Week 16
- H<sub>16</sub>: secukinumab 300 mg is not different to placebo regimen with respect to the change from baseline in HAQ-DI<sup>®</sup> at Week 16
- H<sub>17</sub>: secukinumab 150 mg is not different to placebo regimen with respect to the change from baseline in HAQ-DI<sup>®</sup> at Week 16
- H<sub>18</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to the change from baseline in HAQ-DI<sup>®</sup> at Week 16
- H<sub>19</sub>: secukinumab 300 mg is not different to placebo regimen with respect to the change from baseline in DAS28-CRP at Week 16
- H<sub>20</sub>: secukinumab 150 mg is not different to placebo regimen with respect to the change from baseline in DAS28-CRP at Week 16
- H<sub>21</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to the change from baseline in DAS28-CRP at Week 16
- H<sub>22</sub>: secukinumab 300 mg is not different to placebo regimen with respect to presence of enthesitis at Week 16 in the subset of subjects who have enthesitis at BSL
- H<sub>23</sub>: secukinumab 150 mg is not different to placebo regimen with respect to presence of enthesitis at Week 16 in the subset of subjects who have enthesitis at BSL
- H<sub>24</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to presence of enthesitis at Week 16 in the subset of subjects who have enthesitis at BSL
- H<sub>25</sub>: secukinumab 300 mg is not different to placebo regimen with respect to presence of dactylitis at Week 16 in the subset of subjects who have dactylitis at BSL
- H<sub>26</sub>: secukinumab 150 mg is not different to placebo regimen with respect to presence of dactylitis at Week 16 in the subset of subjects who have dactylitis at BSL
- H<sub>27</sub>: secukinumab 150 mg No Load is not different to placebo regimen with respect to presence of dactylitis at Week 16 in the subset of subjects who have dactylitis at BSL

---

The graphical approach of ([Bretz et al 2009](#)) for sequentially reject testing procedures is used to illustrate the testing strategy:

**Figure 2-1** Hierarchical testing strategy

The family-wise error will be set to  $\alpha=5\%$  and it will be controlled with the proposed hierarchical testing strategy. With this approach, the hypotheses will be separated into two families. The first family will include hypotheses of ACR20 at Week 16 and vdH-mTSS at Week 24 (H<sub>1</sub> through H<sub>6</sub>). The Hypotheses of additional signs and symptoms (H<sub>7</sub> to H<sub>24</sub>) will be the second family. The second family hypotheses will be tested only when all hypotheses in the first family have been rejected.

Each of the hypotheses (H<sub>1</sub>, H<sub>2</sub> and H<sub>3</sub>) for the primary objective (ACR20 at Week 16) for each secukinumab regimen vs. placebo will be tested simultaneously at  $\alpha/3$ . Then based on the rejection of one or more (of H<sub>1</sub>, H<sub>2</sub> and H<sub>3</sub>), the vdH-mTSS at Week 24 endpoint will be tested hierarchically for each dose (through H<sub>4</sub>, H<sub>5</sub> or H<sub>6</sub>) at the level of  $\alpha/3$ . If anyone is rejected, then half (1/2) of  $\alpha/3$  will be passed on to each of the other two hypotheses of ACR20 if they were not already rejected.

If  $H_1$  to  $H_6$  in the first family are all rejected, then AIN457 300 mg for all variables in the second family will be tested hierarchically at  $\alpha$  starting from PASI75 ( $H_7$ ). If all variables in AIN457 300 mg group are rejected then test AIN457 150 mg group and finally AIN 457 150 mg No Load group for all variables in hierarchy. Of note, in the description above, rejection of a hypothesis refers to rejection of the two-sided hypothesis; however the level of a rejected hypothesis is only passed on according to the graphical procedure for the test of another hypothesis if the treatment effect is in favor of secukinumab.

### **2.6.2.2 Model and method of analysis**

#### **ACR50**

Response at Week 16 to ACR50 in the FAS will be evaluated using a logistic regression model with treatment and randomization stratum (TNF $\alpha$  status – naïve or IR) as factors and weight as a covariate.

The primary estimand of ACR50 response is the odds-ratio (each secukinumab dose vs. placebo) at week 16 in the FAS population that would be observed if all patients dropping out or being unblinded before week 16 or having with missing ACR50 data at week 16 were non-responders.

#### **Joint/bone structural damage (vdH-mTSS)**

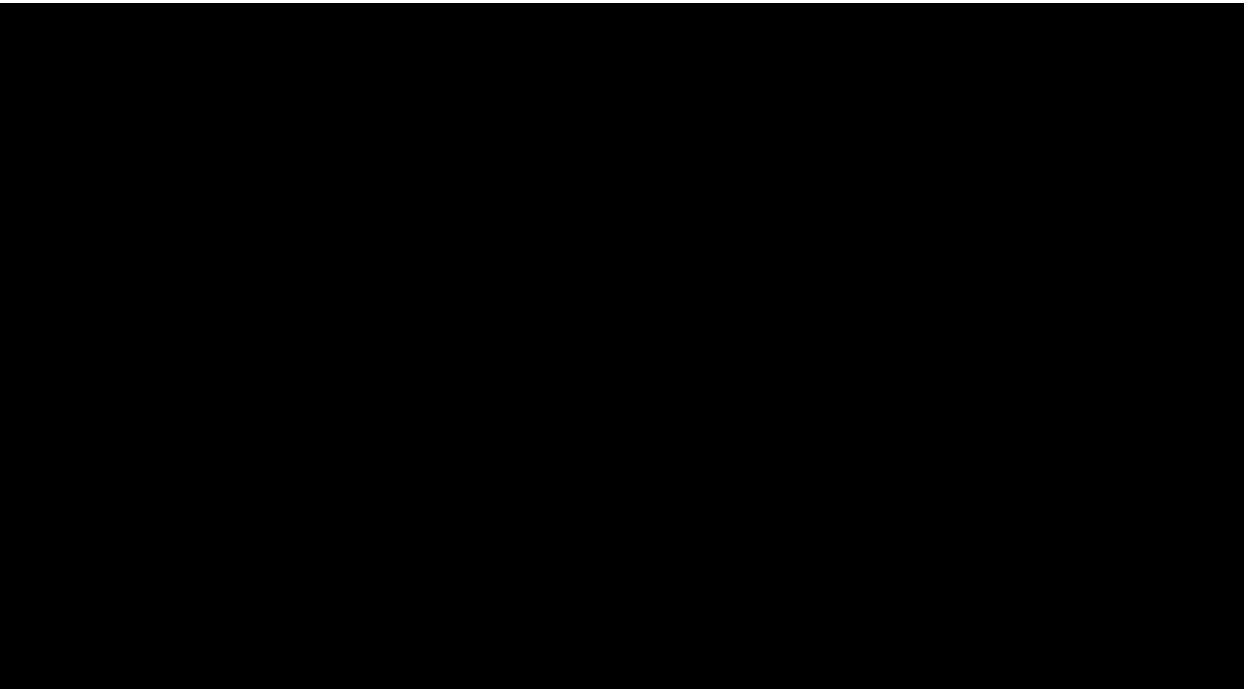
The change at Week 24 from BSL van der Heijde modified Total Sharp Score (vdH-mTSS) will be evaluated using a non-parametric ANCOVA model with treatment regimen and randomization stratum (TNF $\alpha$  status – naïve or IR) as factors, and weight and BSL van der Heijde total modified Sharp score as covariates.

For subjects either on placebo or active, who have a missing vdH-mTSS score at Week 24, linear extrapolation will be used to impute the value at Week 24 if baseline and Week 16 assessments exist. For details about the derivation of the change from baseline see [Section 2.12.4.1](#).

For placebo subjects who meet the criteria for early escape at Week 16, they will be treated as missing for placebo treatment and linear extrapolation will be used to impute the value at Week 24. If BSL or all post- BSL vdH-mTSS values is/are missing for a subject, the subject will be excluded from the analyses.

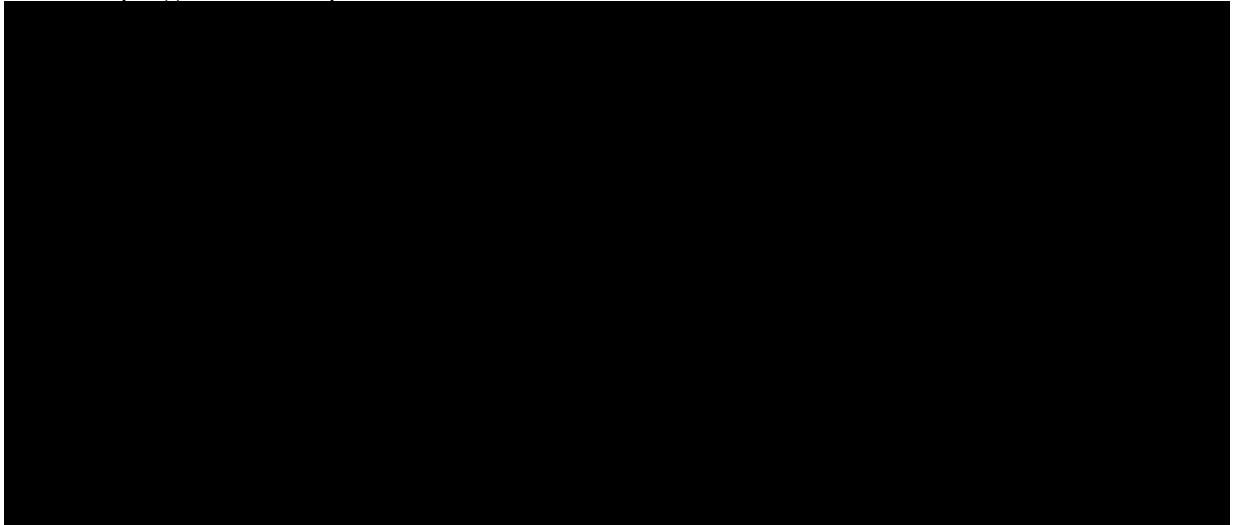
Thus the primary estimand obtained is the difference in the vdH-mTSS scores (secukinumab dose – placebo) at week 24 in the FAS population with at least one baseline and post-baseline vdH-mTSS assessment that would be observed if the possibility for early escape at week 16 were not available to the placebo patient and patients with missing week 24 assessments experienced the same rate of disease progression as up to week 16.

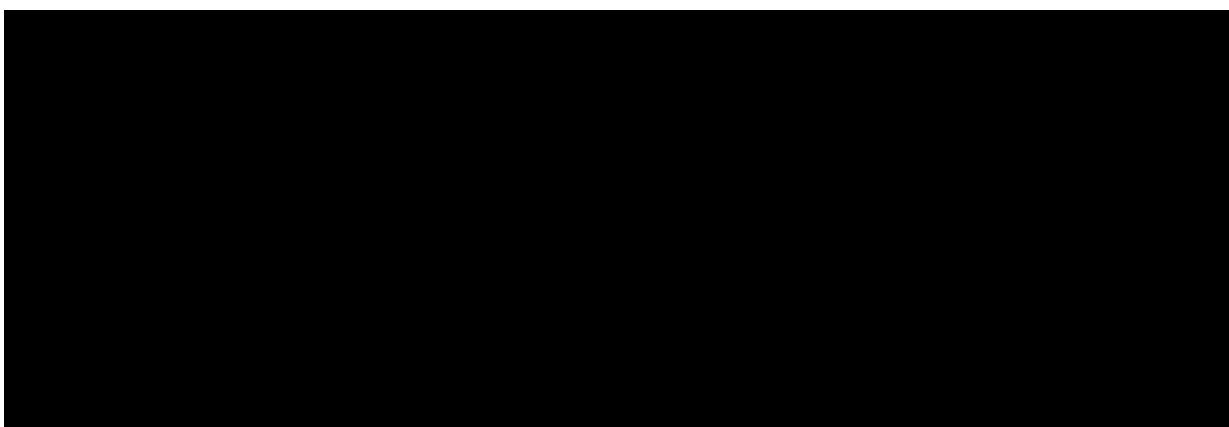




- In addition the binary endpoint of no structural progression will be analysed to assess robustness of the results obtained with the primary estimand of the vdH-mTSS. After applying the linear extrapolation as described above no structural progression is defined as those subjects who have a change in vdH-mTSS at Week 24 relative to baseline  $\leq 0.5$ . The proportion of subjects without disease progression at Week 24 will be evaluated using a logistic regression model with treatment group and randomization strata (TNF $\alpha$  status – naïve or IR) as factors, weight and baseline vdH-mTSS as covariates.

The primary estimand for no structural progression is the odds ratio (secukinumab dose vs. placebo) of no structural progression rates at week 24 in the FAS population of patients with at least one baseline and post-baseline vdH-mTSS assessment that would be observed if the possibility for early escape at week 16 were not available to the placebo patient and patients with missing week 24 assessments experienced the same rate of disease progression as up to week 16.





### **Physical function (HAQ-DI<sup>©</sup>)**

Between-treatment differences in the change in HAQ-DI<sup>©</sup> will be evaluated using a MMRM with treatment, analysis visit and TNF-alpha inhibitor status as factors, and weight and BSL HAQ-DI<sup>©</sup> score as continuous covariates. Treatment by analysis visit and BSL by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for this model. The significance of the treatment effects for secukinumab regimens at different analysis visits will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at the appropriate analysis visits.

The primary estimand for HAQ-DI is the estimated treatment differences (each secukinumab dose vs. placebo) at week 16 in the FAS population applying the MMRM model described above under the MAR assumption for any missing data.

### **Dactylitis**

Presence of dactylitis at Week 16 in the subset of subjects who have dactylitis at BSL will be evaluated using a logistic regression model with treatment and randomization stratum (TNF $\alpha$  status –naïve or IR) as factors and weight as a covariate.

The primary estimand for the presence of Dactylitis is the odds-ratio (each secukinumab dose vs. placebo) at week 16 in the FAS population that would be observed if all patients dropping out or being unblinded before week 16 or having with missing dactylitis data at week 16 were non-responders.

### **Enthesitis**

Presence of enthesitis at Week 16 in the subset of subjects who have enthesitis at BSL will be evaluated using a logistic regression model with treatment and randomization stratum (TNF $\alpha$  status –naïve or IR) as factors and weight as a covariate.

The primary estimand for the present of enthesitis is the odds-ratio (each secukinumab dose vs. placebo) at week 16 in the FAS population that would be observed if all patients dropping out or being unblinded before week 16 or having with missing ethesitis data at week 16 were non-responders.

## **PASI75 and PASI90 response**

PASI75 response and PASI90 at Week 16 will be evaluated for those subjects in whom the assessment occurred due to sufficient skin involvement (at least 3% BSA affected with psoriasis) (which is planned to be a subset of the FAS). These binary variables will be evaluated in the same fashion as ACR response, i.e. a logistic regression model with treatment and randomization strata as factors and weight as a covariate.

The primary estimand of PASI 75 and 90 response is the odds-ratio (each secukinumab dose vs. placebo) at week 16 in the FAS population that would be observed if all patients dropping out or being unblinded before week 16 or having with missing PASI data at week 16 were non-responders.

## **Changes in DAS28-CRP**

Between-treatment differences in the change from baseline in DAS28-CRP will be compared by means of a MMRM with treatment regimen, analysis visit, and TNF-alpha inhibitor status as factors, and weight and BSL as continuous covariates. Treatment by analysis visit and BSL by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for this model. The significance of the treatment effects for secukinumab regimens at different analysis visits will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at the appropriate analysis visits.

The primary estimand for DAS28-CRP is the estimated treatment differences (each secukinumab dose vs. placebo) at week 16 in the FAS population applying the MMRM model described above under the MAR assumption for any missing data.

### **2.6.3 Handling of missing values/censoring/discontinuations**

#### **Missing data**

Missing data for binary efficacy variables for data up to Week 52 will be handled as follows:

- Subjects who drop out of the trial for any reason will be considered non-responders from the time they drop out through Week 52 for active groups or through Week 24 for placebo group.
- Subjects who do not have the required data at baseline and at the specific time point to compute response will be classified as non-responders.

Continuous variables (e.g. ACR20 components) will be analyzed using a mixed-effects repeated measures model (MMRM) which is valid under the missing at random (MAR) assumption. As such, single-point imputation of missing data will not be performed (e.g. LOCF). For analyses of these parameters, if all post-baseline values are missing then these missing values will not be imputed and this subject will be removed from the analysis of the corresponding variable, i.e. it might be that the number of subjects providing data to an analysis is smaller than the number of subjects in the FAS.

## **Premature unblinding**

Patients who were unblinded prior to the scheduled timepoint will be considered non-responders from the time of unblinding up to the end of the placebo-controlled period (Week 24).

## **Rescued patients**

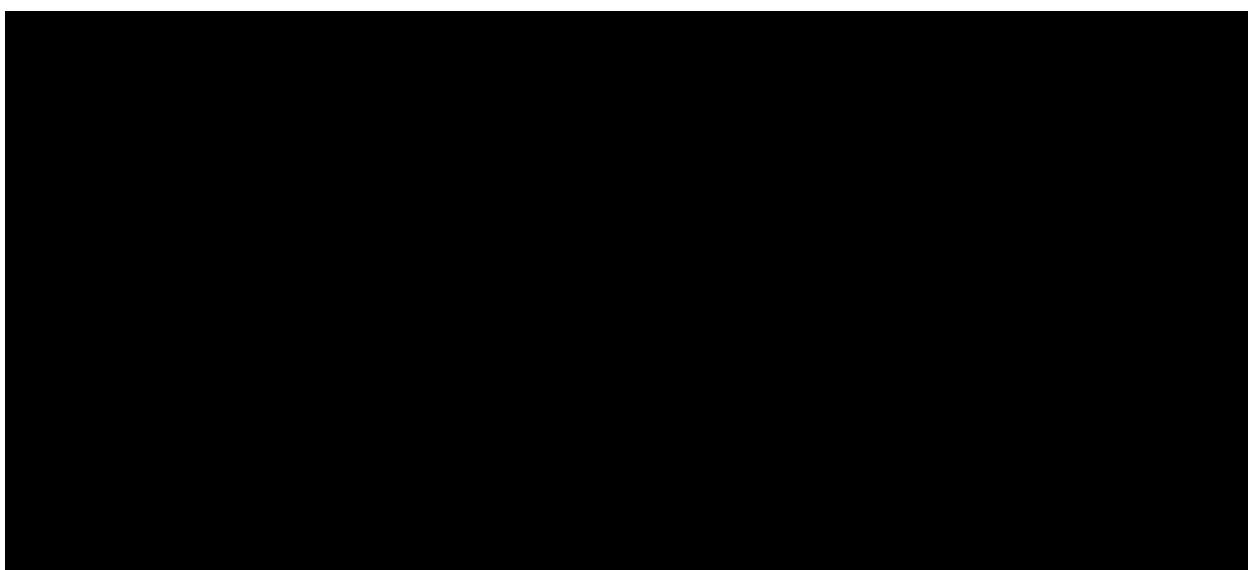
In general if not stated otherwise (as e.g. for vdH-mTSS), the handling of data for subjects who are **rescued at Week 16** will be handled in the following fashion for comparison with placebo at Week 20 and Week 24:

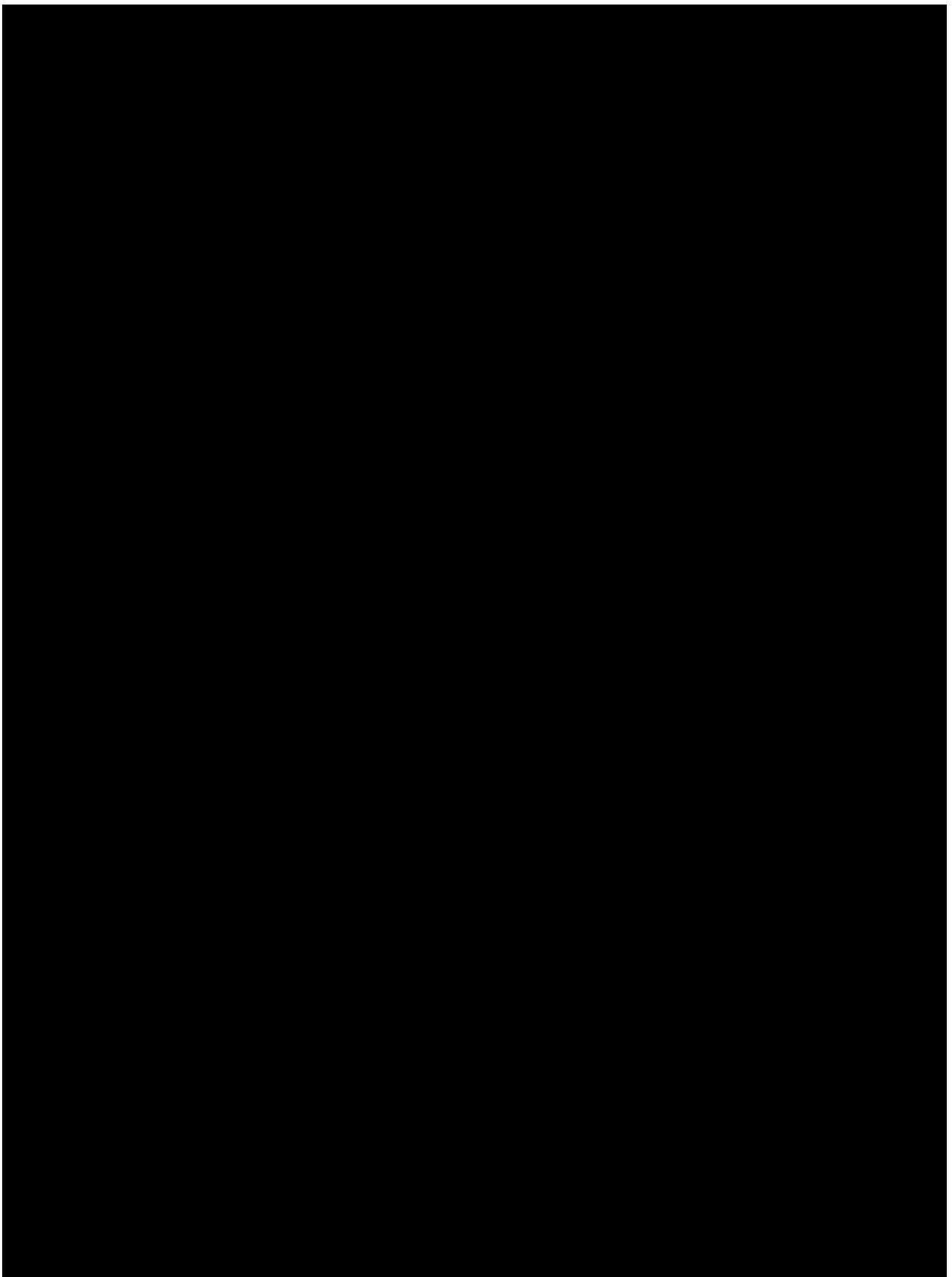
For each binary endpoint, a subject will be considered as a non-responder at Week 20/24 for that endpoint if the subject is a non-responder (including missing imputed non-responder) at Week 16. If a subject is a responder at Week 16 for that endpoint, the actual value at Week 20/24 will be used (missing is considered as non-responder). This will be done for all treatment regimens in order to minimize bias.

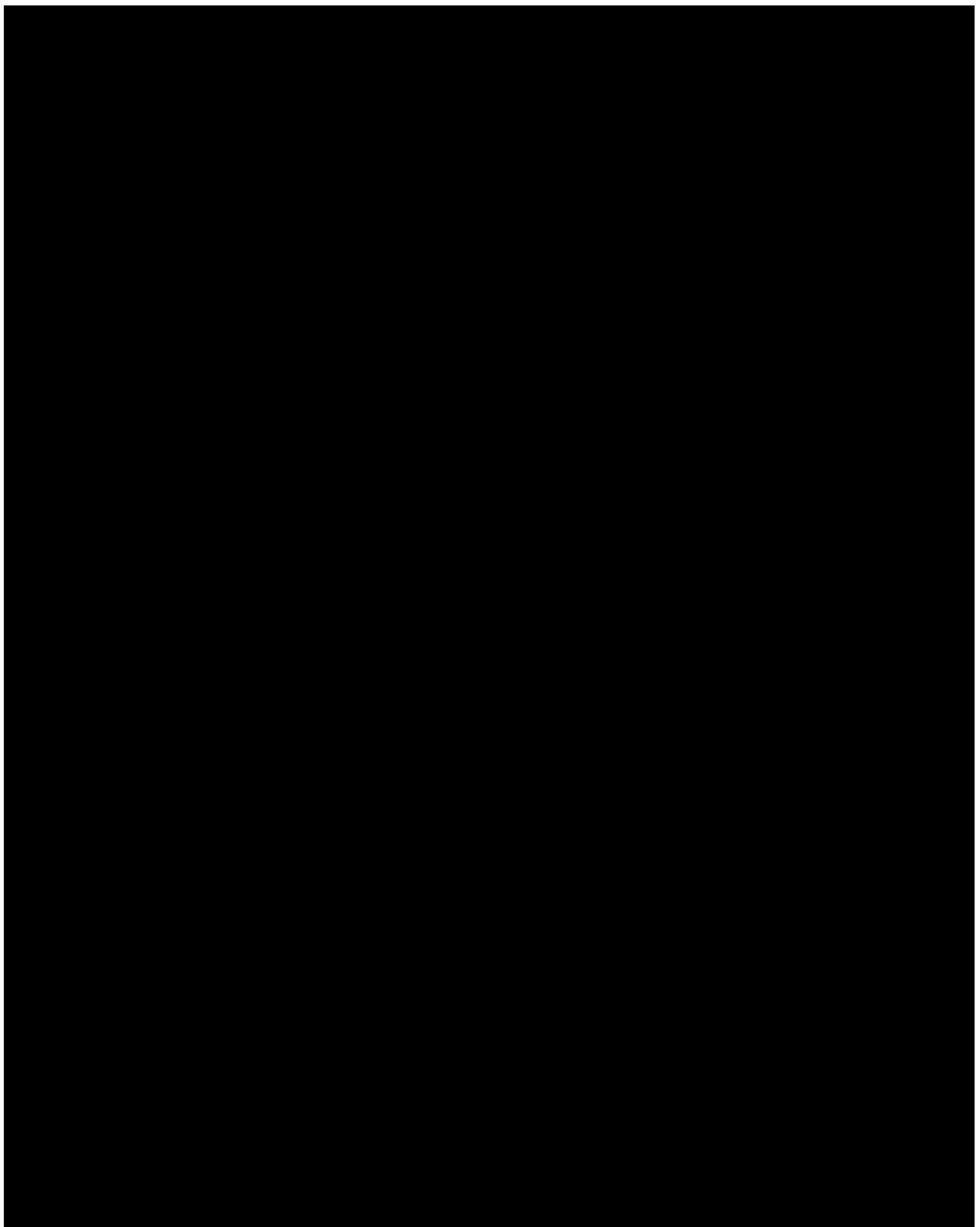
- For continuous endpoints, the goal of the analyses would be to estimate what would have happened if the patients had stayed on the original treatment. Thus, the data collected after a Placebo patient switches to secukinumab will be treated as missing for the placebo group and will be analyzed using a mixed-effects repeated measures model (MMRM) which is valid under the missing at random (MAR) assumption. For secukinumab patients that are rescued, the actual values will be used in the analysis.

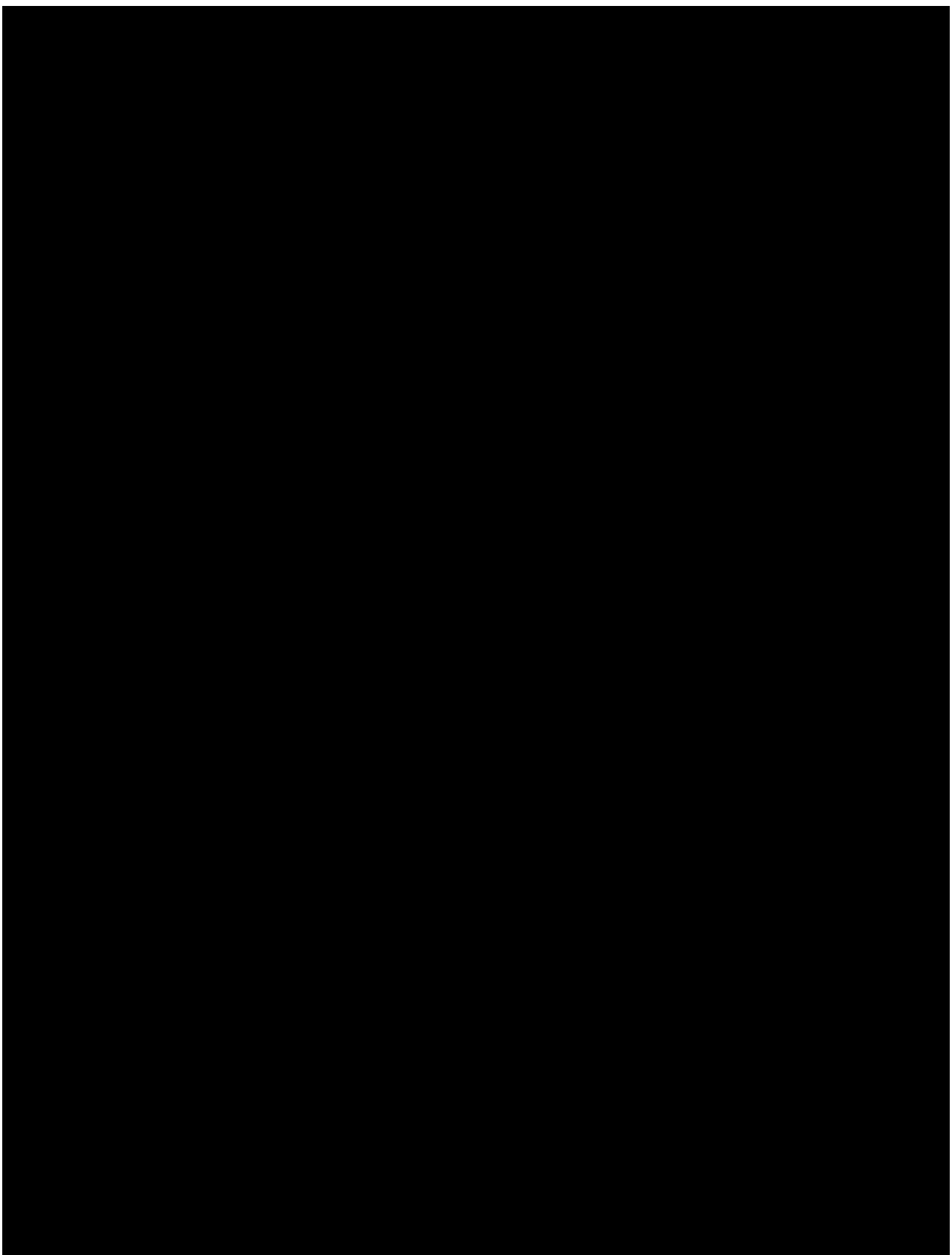
Data collected **after Week 24** will generally be presented as ‘observed case’; i.e. all available data for each time point will be included in the analyses. In addition, multiple imputation will be used for long-term (e.g. after week 52) comparison for selected endpoints.

For comparisons between the active arms the rescue penalty will not be applied as the effect of rescuing is the same for both treatment groups (i.e. the patient continues to receive secukinumab 150 mg or 300 mg every 4 weeks).









## **2.8 Safety analyses**

Summaries may be performed separately for initial (Week 1-16) period and entire treatment period (including follow-up). Week 16 is chosen due to the fact that placebo patients may be rescued as early as week 16. Use of data up to and including the last visit before rescue provides an unbiased comparison between AIN and placebo; data collected beyond week 16 are included in analyses which summarize the entire treatment period.

Safety analyses will be performed on treatment received (or actual treatment) as described below:

The actual treatment or treatment received for summaries of safety data will differ to the treatment assigned at randomization only if a subject received the wrong treatment during the entire study.

For those subjects who received not the treatment randomized, i.e. who received erroneously the wrong treatment at least once, an additional AE listing will be prepared displaying which events occurred after the treatment errors.

### **2.8.1 Adverse events (AEs)**

The crude incidence of treatment emergent adverse events (AEs) (i.e. events started after the first dose of study treatment or events present prior to the first dose of study treatment but increased in severity based on preferred term and on or before last dose + 84 days) will be summarized by primary system organ class and preferred term. Confidence intervals for the crude rate will be derived as described in [Section 4.2.4.1](#) In addition, exposure time-adjusted rates (incidence rate) including 95% confidence intervals will be provided for the entire treatment period (see [Section 4.2.5.1](#)) to adjust for differences in exposure. Graphical displays of the crude incidence rates and exposure-adjusted rates will be presented for all AEs and serious AEs by system organ class.

Adverse events reported will be presented in descending frequency according to its incidence in total secukinumab group (combining all secukinumab treatment groups) starting from the most common event. Summaries (crude incidences only) will also be presented for AEs by severity and for study treatment related AEs. If a particular AE 'severity' is missing, this variable will be listed as missing and treated as missing in summaries. If a subject reported more than one adverse event with the same preferred term, the adverse event with the greatest severity will be presented. If a subject reported more than one adverse event within the same primary system organ class, the subject will be counted only once with the greatest severity at the system organ class level, where applicable.

Separate summaries will be provided for adverse events suspected to be related to study drug, deaths, serious adverse events, and adverse events leading to discontinuation and adverse events leading to temporary dose interruption.

Adverse events will also be reported separately by SMQ according to MedDRA. The MedDRA version used for reporting the study will be described in a footnote.

Non-treatment emergent adverse events will be listed.

Algorithms for date imputations will be provided in RAP M8.

For serious adverse events (SAEs) occurred during screening a listing will be prepared for all subjects screened including screening failures.

When adjudication is required of major cardiovascular events, a listing of those types of events as reported by the investigator and confirmed by adjudication will be provided.

The safety analyses that will be performed for treatment emergent AEs and on treatment labs, ECG and vital signs for each analysis period is described in [Table 2-1](#).

**Table 2-1      Overview of analyses on some safety endpoints**

Analysis period	AEs & SAEs	AEs by severity	Study drug related AEs	AEs-SMQ	Risk	Notables for (vitals/ECG), lab criteria
Day 1 – Week 8	crude incidence					
Day 1 – Week 16	crude incidence	crude incidence	crude incidence	crude incidence	crude incidence	crude incidence
Entire Treatment	crude incidence	crude incidence	crude incidence	exposure time	crude incidence	crude incidence
	exposure time			adjusted incidence	exposure time	
	adjusted incidence*				adjusted incidence	

\*Exposure-adjusted incidence rates will be done for the following :

at the PSOC for AE and SAE and Level 1 for Risks and SMQ analyses

at the PT level for common AEs, which is defined as at least 2% of the patients in the combined AIN457 groups during the initial treatment period (i.e. up to week 24) or events that had an incidence rate of at least 5.0 cases per 100 subject-years in the combined AIN457 groups during the entire treatment period

## 2.8.2      Laboratory data

The summary of lab data will only include on treatment data, which are defined as those lab assessments after the first dose of study treatment and on or before last dose + 84 days.

The summary of laboratory evaluations will be presented for three groups of laboratory tests (hematology, chemistry and urinalysis). In addition to the individual laboratory parameters the ratios “total cholesterol / HDL” and “apolipoprotein B / apolipoprotein A1” will be derived and summarized.

For urinalysis, frequency tables will be presented.

Descriptive summary statistics for the change from baseline to each study visit will be presented. These descriptive summaries will be presented by laboratory test and treatment

group. Change from baseline will only be summarized for subjects with both baseline and post baseline values and will be calculated as:

$$\text{change from baseline} = \text{post baseline value} - \text{baseline value}$$

For each parameter, the maximum change (maximum decrease and maximum increase) from baseline, will be analyzed analogously.

In addition, shift tables will be provided for all parameters to compare a subject's baseline laboratory evaluation relative to the visit's observed value. For the shift tables, the normal laboratory ranges will be used to evaluate whether a particular laboratory test value was normal, low, or high for each visit value relative to whether or not the baseline value was normal, low, or high. If appropriate, the shifts to the most extreme laboratory test value within a treatment phase (either initial up to week 24 or entire) will be presented as well (including category "high and low"). These summaries will be presented by laboratory test and treatment group.

The following laboratory parameters will be analyzed with respect to numerical Common Terminology Criteria for Adverse Events (CTCAE) grades, given in [Table 2-2](#): hemoglobin, platelets, white blood cell count, neutrophils, lymphocytes, creatinine, total bilirubin (TBL), gamma-glutamyl transferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), glucose, cholesterol, triglycerides (TG).

These summaries will be split into hematology and chemistry.

**Table 2-2 CTCAE grades for laboratory parameters to be analyzed**

CTCAE v4.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
HGB decreased (Anemia)	<LLN - 100 g/L	<100 - 80 g/L	<80 g/L	
Platelet count decreased	<LLN - 75.0 x10e9 /L	<75.0 - 50.0 x10e9 /L	<50.0 - 25.0 x10e9 /L	<25.0 x 10e9 /L
White blood cell decreased	<LLN - 3.0 x 10e9 /L	<3.0 - 2.0 x 10e9 /L	<2.0 - 1.0 x 10e9 /L	<1.0 x 10e9 /L
Neutrophil count decreased	<LLN - 1.5 x 10e9 /L	<1.5 - 1.0 x 10e9 /L	<1.0 - 0.5 x 10e9 /L	<0.5 x 10e9 /L
Lymphocyte count decreased	<LLN - 0.8 x 10e9/L	<0.8 - 0.5 x 10e9 /L	<0.5 - 0.2 x 10e9 /L	<0.2 x 10e9 /L
Creatinine increased*	>1 - 1.5 x baseline; >ULN - 1.5 x ULN	>1.5 - 3.0 x baseline; >3.0 x ULN	>3.0 baseline; >6.0 x ULN	
TBL increased	>ULN - 1.5 x ULN	>1.5 - 3.0 x ULN	>3.0 - 10.0 x ULN	>10.0 x ULN
GGT increased	>ULN - 2.5 x ULN	>2.5 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
ALT increased	>ULN - 3.0 x ULN	>3.0 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
AST increased	>ULN - 3.0 x ULN	>3.0 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
ALP increased	>ULN - 2.5 x ULN	>2.5 - 5.0 x ULN	>5.0 - 20.0 x ULN	>20.0 x ULN
Glucose increased (Hyperglycemia )	>ULN - 8.9 mmol/L	>8.9 - 13.9 mmol/L	>13.9 - 27.8 mmol/L	>27.8 mmol/L

CTCAE v4.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
Glucose decreased (Hypoglycemia)	<LLN - 3.0 mmol/L >ULN - 7.75 mmol/L	<3.0 - 2.2 mmol/L >7.75 - 10.34 mmol/L	<2.2 - 1.7 mmol/L >10.34 - 12.92 mmol/L	<1.7 mmol/L >12.92 mmol/L
Cholesterol high	mmol/L	mmol/L	mmol/L	mmol/L
Hypertriglyceride mia	1.71 - 3.42 mmol/L	>3.42 - 5.7mmol/L	>5.7 - 11.4 mmol/L	>11.4 mmol/L

\*Note: for "creatinine increased" the baseline criteria do not apply

Shift tables will be presented comparing baseline laboratory result (CTCAE grade) with the worst results (expressed in CTCAE grade) during the treatment phase (either initial up to week 24 or entire) analyzed. Of note, baseline will be defined as last assessment prior to first dosing in initial treatment phase.

Subjects with abnormal laboratory values will be listed and values outside the normal ranges will be flagged.

Summaries for newly occurring or worsening clinically notable lipid abnormalities will also be provided cumulatively for each of the following parameters and categories:

- HDL:
  - <=LLN
  - <0.8 x LLN
- LDL, cholesterol, triglycerides:
  - >=ULN
  - >1.5 x ULN
  - >2.5 x ULN

Newly occurring or worsening liver enzyme abnormalities will also be summarized based on the event criteria given in [Table 2-3](#):

**Table 2-3 Liver-related events**

Parameter	Criterion
ALT	>3xULN; >5xULN; >8xULN; >10xULN, >20xULN
AST	>3xULN; >5xULN; >8xULN >10xULN; >20xULN
ALT or AST	>3xULN; >5xULN; >8xULN >10xULN; >20xULN
TBL	>1.5xULN, >2xULN, >3xULN,
ALP	>2xULN, >3xULN. >5xULN
ALT or AST & TBL	ALT or AST>3xULN & TBL >2xULN; ALT or AST >5xULN & TBL >2xULN; ALT or AST >8xULN & TBL >2xULN; ALT or AST >10xULN & TBL >2xULN
ALP & TBL	ALP >3xULN & TBL >2xULN ALP >5xULN & TBL >2xULN
ALT or AST & TBL & ALP	ALT or AST>3xULN & TBL >2xULN & ALP <2xULN ( <b>Hy's Law</b> )

Parameter	Criterion
	<p>Note: elevated ALP may suggest obstruction as a consequence of gall bladder or bile duct disease; ALP may also be increased in malignancy. FDA therefore terms Hy's Law cases as indicators of <i>pure hepatocellular injury</i>. This does not mean that cases of ALT or AST &gt;3xULN &amp; TBL &gt;2xULN &amp; ALP <math>\geq</math>2xULN may not result in severe DILI.</p>

Notes:

- 1) In studies which enroll subjects with pre-existing liver disease, baseline LFT may be increased above ULN; in such a case it is meaningful to add the condition "and worse than baseline" to the abnormality criteria

For a combined criterion to be fulfilled, all conditions have to be fulfilled on the same visit. The criteria are not mutually exclusive, e.g. a subject with ALT = 6.42xULN is counted for ALT >3xULN and ALT>5x ULN.

Individual subject data listings will be provided for subjects with abnormal laboratory data. Data of subjects with newly occurring or worsening liver enzyme abnormalities will be listed in an additional listing.

Box plots over time will be presented for selected laboratory parameters (neutrophils, liver and lipid parameters).

In addition, individual subject data listings will be provided for all parameters.

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

### 2.8.3 Vital signs

The summary of vital signs will only include on treatment data, which are defined as those vital sign measurements after the first dose of study treatment and on or before last dose + 84 days.

Analysis in vital sign measurement using descriptive summary statistics for the change from baseline for each post-baseline visit will be performed. These descriptive summaries will be presented by vital sign and treatment group. Change from baseline will only be summarized for subjects with both baseline and post-baseline values and will be calculated as:

change from baseline = post-baseline value – baseline value

The number and percentage of subjects with newly occurring notable vital signs will be presented. Criteria for notable vital sign abnormalities are provided in [Table 2-4](#) below.

**Table 2-4 Criteria for notable vital sign abnormalities**

Vital sign (unit)	Notable abnormalities
Systolic blood pressure (mmHg)	$\geq 140$ mmHg or $< 90$ mmHg
Diastolic blood pressure (mmHg)	$\geq 90$ mmHg or $< 60$ mmHg
Pulse (bpm)	$> 100$ bpm or $< 60$ bpm

## **2.8.4      Electrocardiogram (ECG)**

The summary of ECG will only include on treatment data, which are defined as those ECG measurements after the first dose of study treatment and on or before last dose + 84 days.

The following quantitative variables will be summarized: ventricular rate, RR interval, PR interval, QRS duration, QT interval, and corrected QT interval (QTc). Only Fridericia (QTcF) corrections will be presented for QTc.

QTc will be summarized by computing the number and percentage of subjects (including 95% confidence intervals for pooled analyses, e.g. SCS) with:

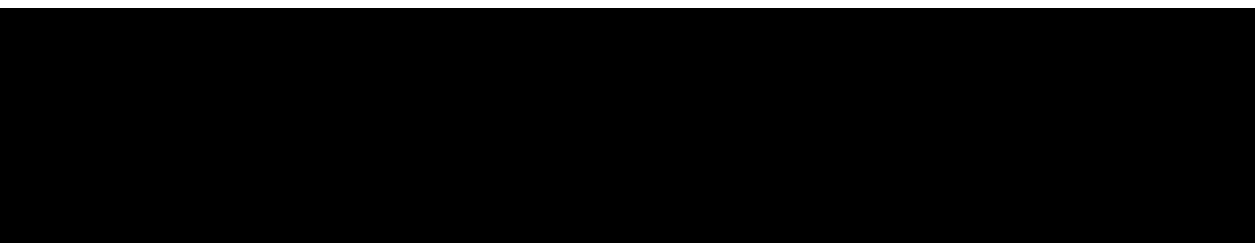
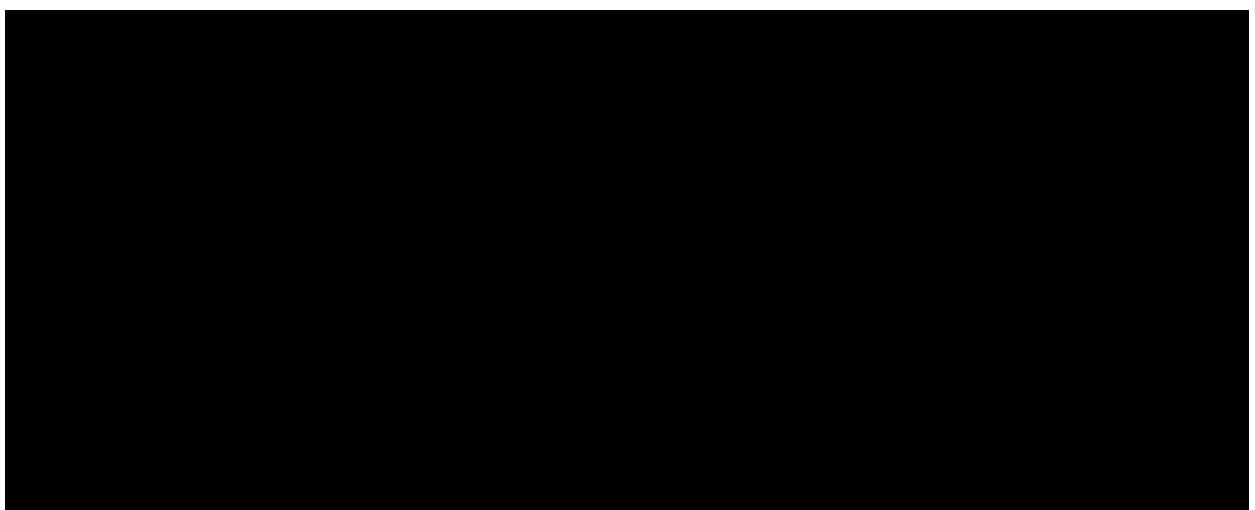
- QTcF > 500 msec
- QTcF > 480 msec
- QTcF > 450 msec
- QTcF changes from baseline > 30 msec
- QTcF changes from baseline > 60 msec

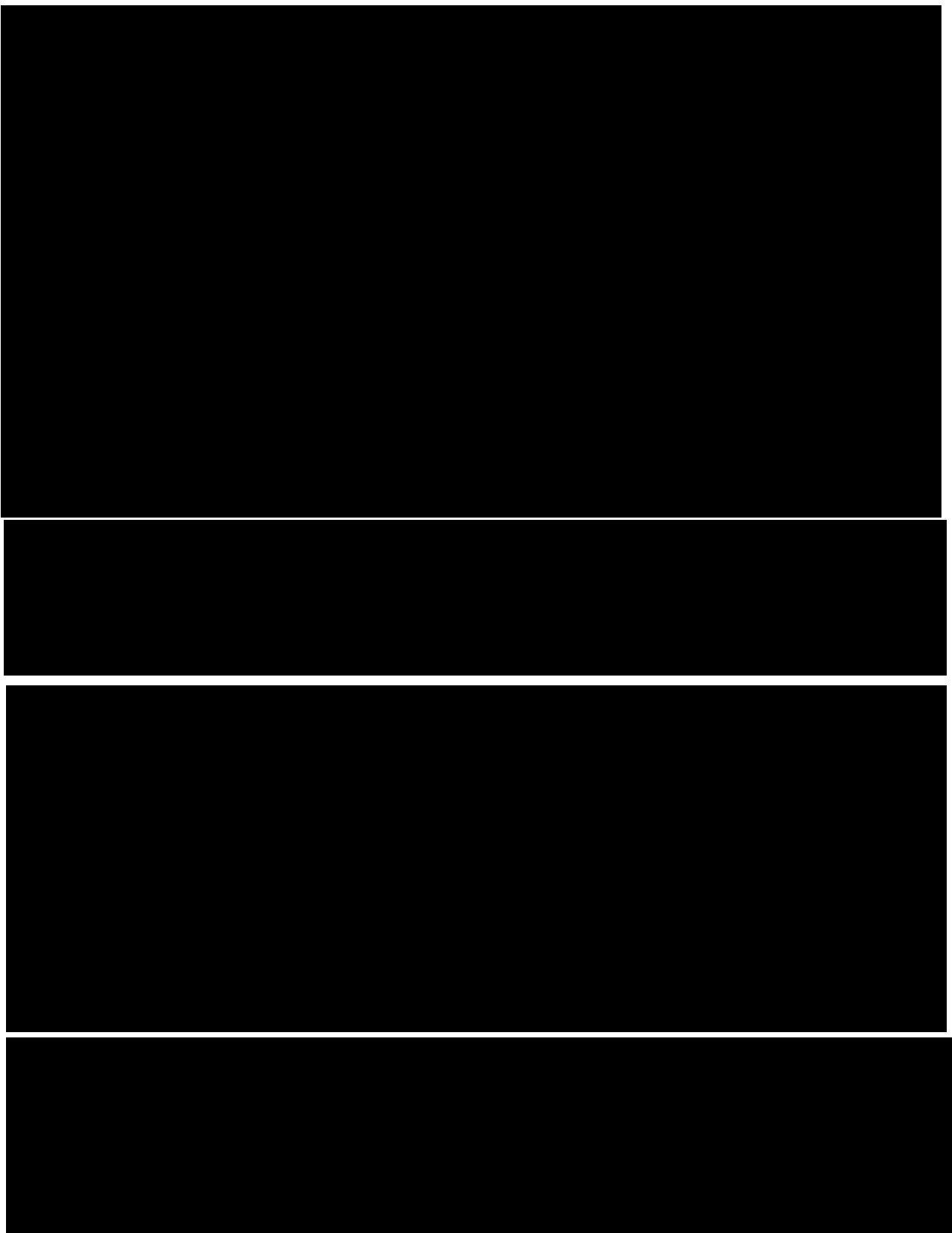
PR > 250 msec

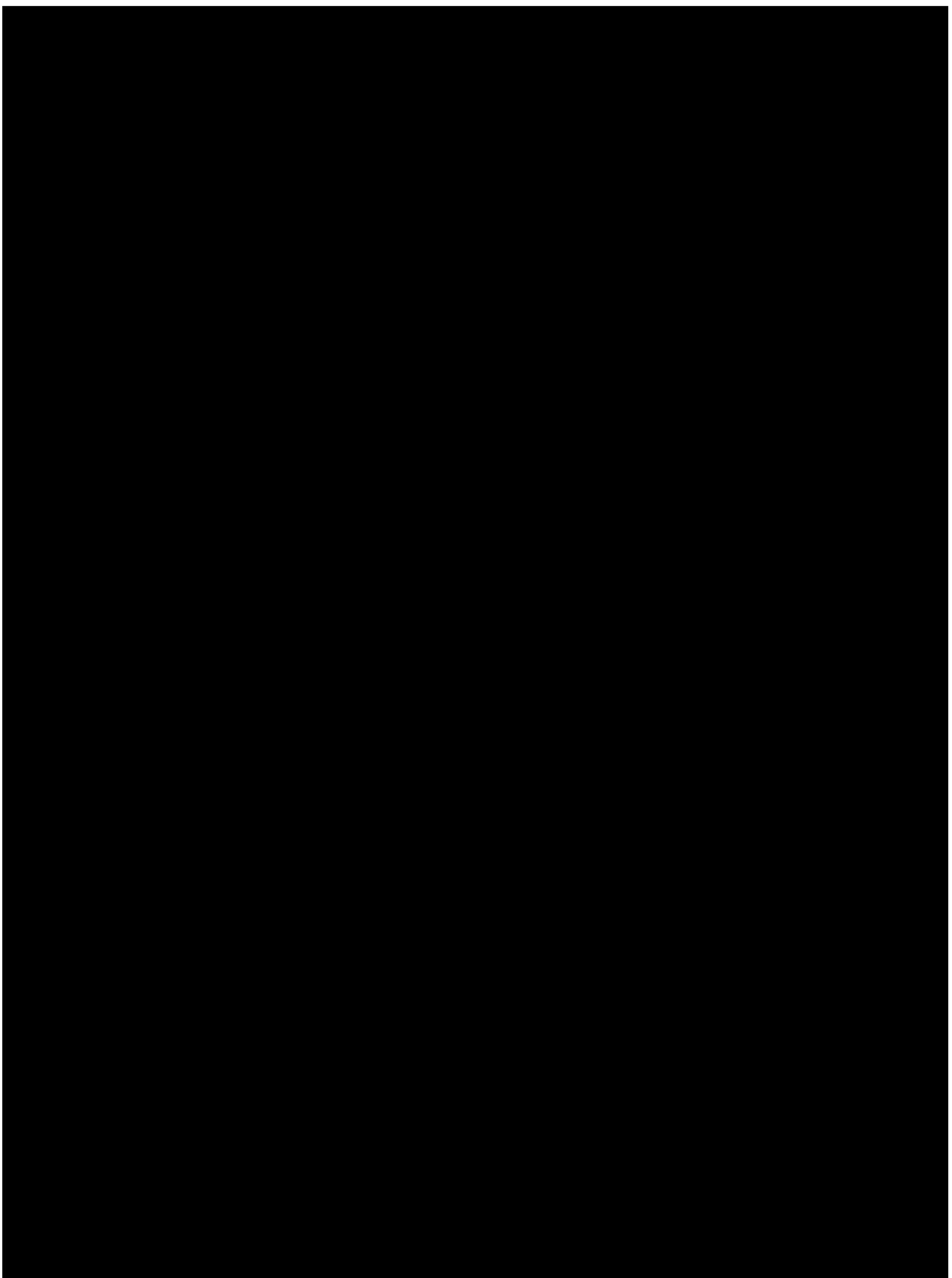
Summary statistics will be presented for ECG variables by visit and treatment group.

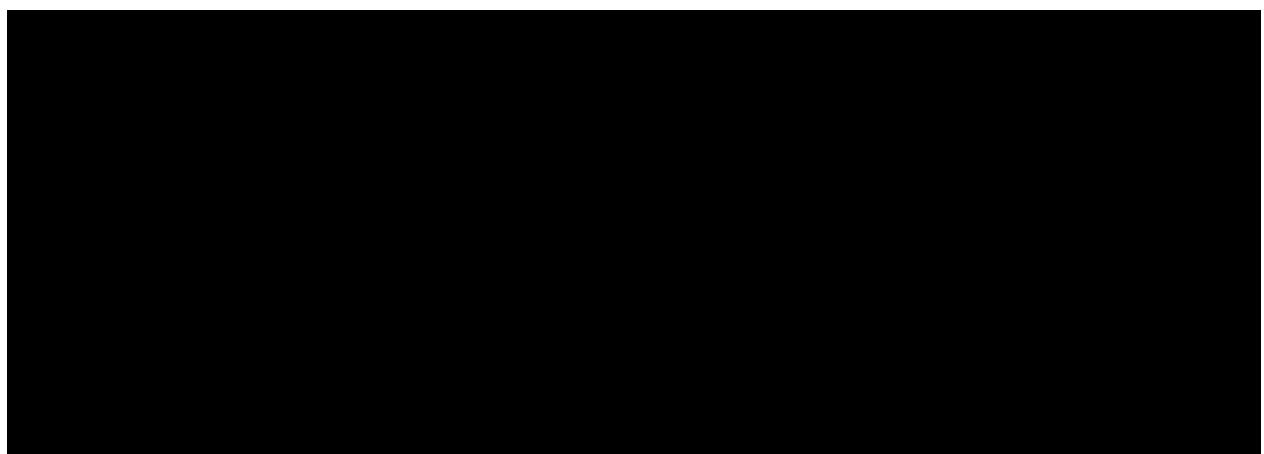
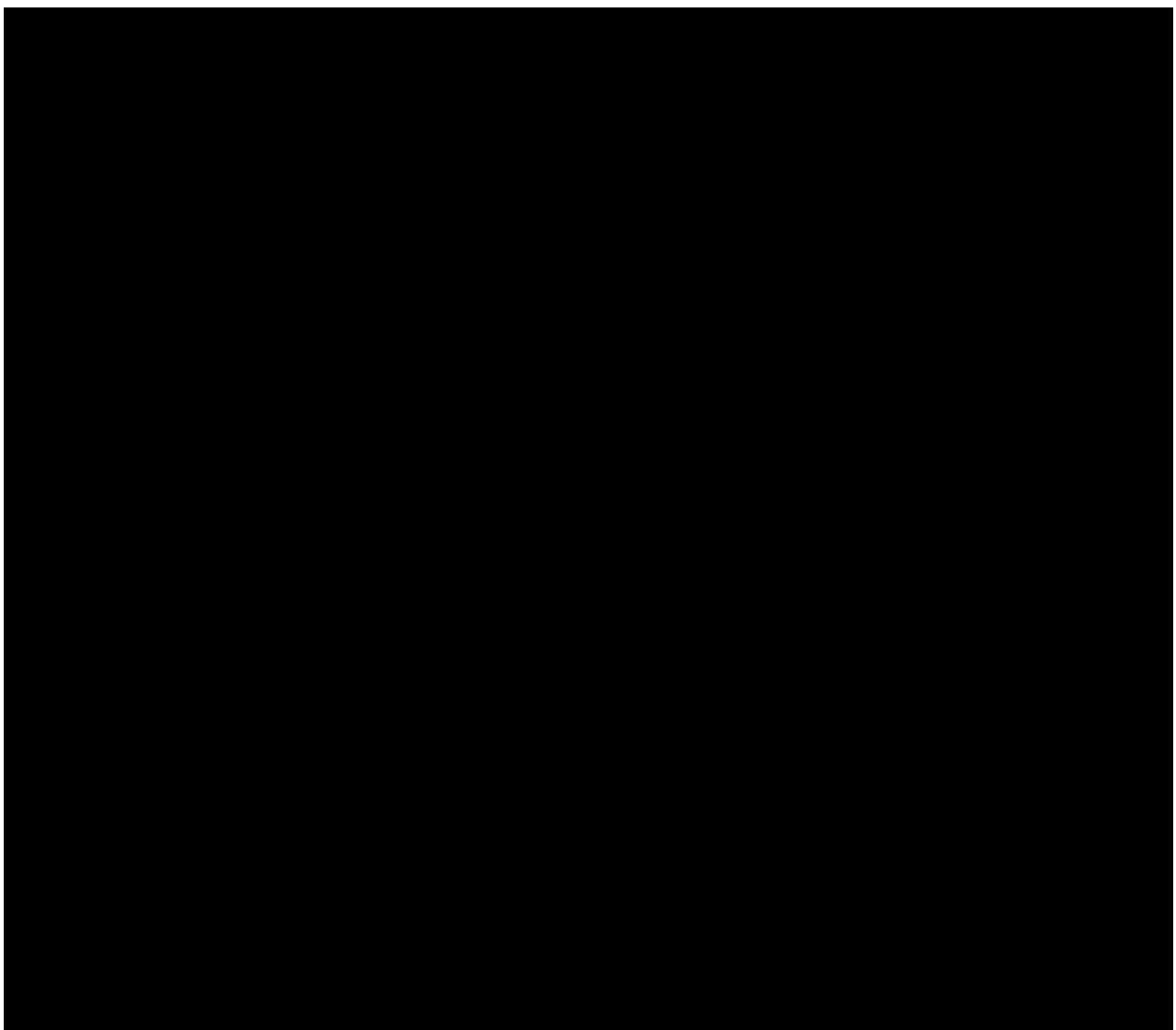
In addition, shift tables comparing baseline ECG interpretation (normal, abnormal, not available, total) with the worst on-study interpretation (normal, abnormal, not available, total) will be provided.

A listing of all newly occurring or worsening abnormalities will be provided, as well as a by-subject listing of all quantitative ECG parameters.









## 2.12.4 Efficacy Evaluation

### 2.12.4.1 Description of efficacy variables

#### ACR 20/50/70

ACR20 is a binary response variable defined for each subject. A subject will be considered a responder according to ACR20 criteria if he/she has at least (i.e.,  $\geq$ ):

- 20% improvement from baseline in tender 78-joint count
- 20% improvement from baseline in swollen 76-joint count
- 20% improvement from baseline in at least 3 of the following 5 measures:
  - Patient's assessment of PsA pain (VAS 100 mm)
  - Patient's global assessment of disease activity (VAS 100 mm)
  - Physician's global assessment of PsA disease activity (VAS 100 mm)
  - Patient self-assessed disability (Health Assessment Questionnaire [HAQ<sup>©</sup>] score)
  - Acute phase reactant (C-reactive protein [hsCRP]) or Erythrocyte sedimentation rate (ESR).

In the definition above, the *baseline* value refers to the last measurement made prior to administration of the first dose of study treatment.

The primary endpoint is the proportion of subjects achieving ACR20 at Week 24. Primarily, CRP will be used to calculate ACR response: ESR will only be used in the event CRP is missing.

ACR50 and ACR70 are defined in the same way as ACR20 by replacing the 20% with 50% and 70% improvement from baseline, respectively.

ACRn represents the percent improvement on the continuous scale and from ACRn one can directly calculate ACR20, ACR50, and ACR70 using the appropriate cutoffs. This variable is defined as:

ACRn =  $\min(x_1, x_2, x_3)$ , where

- $x_1$  = % improvement from baseline in tender 78-joint count
- $x_2$  = % improvement from baseline in swollen 76-joint count

and  $x_3$  = 3<sup>rd</sup> largest value of  $x_4, x_5, x_6, x_7, x_8$  where,

- $x_4$  = % improvement from baseline in Patient's assessment of PsA pain (VAS 100 mm)
- $x_5$  = % improvement from baseline in Patient's global assessment of disease activity (VAS 100 mm)
- $x_6$  = % improvement from baseline in Physician's global assessment of PsA disease activity (VAS 100 mm)
- $x_7$  = % improvement from baseline in Patient self-assessed disability (Health Assessment Questionnaire [HAQ<sup>©</sup>] score)
- $x_8$  = % improvement from baseline in Acute phase reactant (C-reactive protein [hsCRP]) or Erythrocyte sedimentation rate (ESR)

ACRn can be computed even if up to two values of  $x_4$ ,  $x_5$ ,  $x_6$ ,  $x_7$ ,  $x_8$  are missing. ACRn, theoretically, cannot be computed, if one or both of  $x_1$ ,  $x_2$  is/are missing OR more than three values of  $x_4$ ,  $x_5$ ,  $x_6$ ,  $x_7$ ,  $x_8$  are missing.

### Joint/bone structural damage

The score of primary interest is the total van der Heijde modified Sharpe Score (vdH-mTSS) (van der Heijde 1999), but the erosion score and joint space narrowing score will be analyzed in similar fashion.

#### Derivation of vdH-mTSS

The **total vdH-mTSS** is the sum of scores for erosions and joint space narrowing (JSN) in both hands and feet added together. The erosion and JSN scores are derived as follows:

- **Erosions** will be assessed each hand (20 locations per hand) and each foot (6 locations per foot). The maximum erosion score is 200 for all 40 hand locations, and 120 for all 12 feet locations. Thus, the total possible erosion score is 320. Individual scores range between 0-5 joints of the hands and 0-10 for joints of the feet.
- **Joint space narrowing (JSN)** will be assessed in each hand (20 locations per hand) and each foot (6 locations per foot). The maximum score is 160 for all 40 hand joints, and 48 for all 12 feet joints. Thus, the total possible JSN score is 208. The individual scores of each joint range between 0-4.
- Pencil-in-cup: Osteolysis of the proximal phalanx and the base of the distal phalanx resulting in a pencil like proximal phalanx covered by cup like base of the distal phalanx. Pencil-in-cup will be scored as “P” where applicable.
- Gross Osteolysis: Osteolysis of the phalanx resulting in a loss of the normal joint structure, usually accompanied by shortening of the length of the phalanx. Gross osteolysis will be scored as “G” where applicable.
- If a joint or bone is not visible (e.g. poor film quality, missing imaging, severe misalignment, flexion deformity, dislocation) at the timepoint, the individual joint or bone will be coded as Not Visible (N).
- If radiographs at the timepoint show a joint or bone with surgical fusion, replacement (prosthesis), or amputation, then the joint or bone will be scored Surgically Modified (S).

To obtain the **total vdH-mTSS**, scores for erosions and JSN in both the hands and feet will be added together. Any “P” or “G” will be considered the maximal score for the feature (erosions and JSN) per location in the calculation of the total vdH-S score. Any “N” or “S” will be considered null in the calculation of the total vdH-S score. The range of scores is 0 - 528.

#### Imputation of missing data on joint and segment level:

1. Step - On joint level: The joints are divided into 10 segments according to Table 2-5. In each segment an adequacy threshold is defined. For each segment, when the change from baseline values are available for at least the threshold number of joints, imputation will be applied with the following steps for a given follow-up visit:
  1. The change from baseline will be calculated for the segment and missing joint change from baseline values will be imputed by the within-segment mean change from baseline..

2. Missing baseline and follow-up values will be imputed based on the available data and imputed change from baseline values. If insufficient data is present the mean baseline or mean follow-up value of this segment will be imputed.
3. During the imputation of baseline or post-baseline values implausible values can occur. The following correction step is implemented depending on the value obtained:
  - If a negative value was derived (e.g. for a follow-up visit) this is replaced by 0.
  - If a value greater than the maximum possible score is derived (i.e. >5 for hands and >10 for feet e.g. for the baseline visit) it is replaced by the maximum possible score.

The change from baseline for this joint is updated after this step accordingly.

This algorithm implies that different baseline values can occur for the different follow-up visits if imputation of values is needed.

Otherwise (i.e. if less than threshold number of joints have available change from baseline values), the segment change from baseline will be missing.

2. Step - On segment level: When  $\geq 6$  segments of the 10 segments in total have evaluable change from baseline values, similar as on joint level imputation will be applied with the following steps for a given follow-up visit:
  1. The change from baseline will be calculated across all segments and missing segments change from baseline values will be imputed with the across segments mean change from baseline.
  2. Missing baseline and follow-up values will be imputed based on the available data and imputed change from baseline values. If insufficient data is present the mean baseline or mean follow-up values across segments will be imputed.

Otherwise (i.e. if  $\leq 5$  segments have evaluable change from baseline values), the overall change from baseline value for the segment will be missing.

**Table 2-5 Segmental distribution**

Segment	Total number of joints	Adequacy threshold	Joints at one side	Location of joints
PIP erosions	8	5	PIP2, PIP3, PIP4, PIP5	Hand
DIP erosions	8	5	DIP2, DIP3, DIP4, DIP5	Hand
MCP + thumb erosions	12	7	MCP1, MCP2, MCP3, MCP4, MCP5, INTERPHALANGEAL JOINT OF THE HAND	Hand
Wrist erosions	12	7	FIRST METACARPAL BONE, DISTAL RADIUS, DISTAL ULNA, TRAPEZOID-TRAPEZIUM, NAVICULAR BONE, LUNATE BONE	Hand
Foot erosions	12	7	MTP1, MTP2, MTP3, MTP4, MTP5, INTERPHALANGEAL JOINT 1	Foot
PIP JSN	8	5	PIP2, PIP3, PIP4, PIP5	Hand
DIP JSN	8	5	DIP2, DIP3, DIP4, DIP5	Hand

MCP + thumb JSN	12	7	MCP1, MCP2, MCP3, MCP4, MCP5, Hand INTERPHALANGEAL JOINT OF THE HAND
Wrist JSN	12	7	CMC3, CMC4, CMC5, RADIOCARPAL, SCAPHOID- TRAPEZIUM, CAPITATE- NAVICULAR-LUNATE
Foot JSN	12	7	MTP1, MTP2, MTP3, MTP4, MTP5, Foot INTERPHALANGEAL JOINT 1

### **Handling of multiple X-ray readers**

The readings of the x-rays and the scoring will be performed centrally. Two central independent primary radiograph readers, both blinded to treatment arm and radiograph sequence, will analyze the digitized images. In the case that adjudication is needed, the primary independent reviewers will re-read images first. The re-read results will be considered as the final primary read results. Following the re-reads, adjudication will commence if the adjudication is still needed. A third reader (i.e. an adjudicator different from the primary reviewers for a given subject) will make an independent assessment.

Each reader's total score is calculated separately at the patient level for BSL, Follow Up (FU) and Change from BSL as detailed in [Appendix 5-3](#). The statistical analysis will use the mean total score from all three readers, if available, or the mean total score from the two readers otherwise.

### **Handling of rescued/early escape patients data**

Two scenarios might be observed:

Subjects following the original protocol who are non-responder at Week 16 (not achieving a  $\geq 20\%$  improvement from baseline in both TJC and SJC) will have their hands/wrists and feet X-rays taken at Week 16 and no additional X-ray at week 24.

Subjects following the amendment 1 of the protocol who are non-responders at week 16 are expected to have both a week 16 and week 24 X-ray taken.

All subjects (irrespective of the protocol version) who are responder at Week 16 ( $\geq 20\%$  improvement from baseline in both tender and swollen joint counts) will have their hands/wrists and feet X-rays taken at Week 24.

For patients receiving rescue therapy (who meet the criteria for early escape at Week 16) their week 24 X-ray assessment will be treated as missing for placebo treatment and linear extrapolation will be used to impute the value at Week 24.

### **Missing data following imputation and averaging over readers**

Following application of the imputation algorithm during data derivation, there will be no further imputation of data during the analysis, even if there are missing joint scores e.g. at baseline. If BSL or all post- BSL total modified Sharp score/s is/are missing for a subject, the subject will be excluded from the analyses.

## **Health Assessment Questionnaire - Disability Index (HAQ-DI)**

The Health Assessment Questionnaire (HAQ©) was developed by Stanford University and is one of the most widely used measures to assess the long-term influence of chronic disease on a subject's level of functional ability and activity restriction. The disability assessment component of the HAQ (Health Assessment Questionnaire – Disability Index), the HAQ-DI, assesses a subject's level of functional ability and includes questions of fine movements of the upper extremity, locomotor activities of the lower extremity, and activities that involve both upper and lower extremities. There are 20 questions in eight categories of functioning including dressing, rising, eating, walking, hygiene, reach, grip, and usual activities. The stem of each item asks over the past week "Are you able to ..." perform a particular task. Each item is scored on a 4-point scale from 0 to 3, representing normal (normal, no difficulty [0]), some difficulty [1], much difficulty [2], and unable to do [3].

Scoring for the eight functional categories and overall disability index scoring will be performed as follows:

There are eight categories; first score within each category:

- Dressing and Grooming, includes items 1 and 2
- Arising, includes items 3 and 4
- Eating, includes items 5, 6 and 7
- Walking, includes items 8 and 9
- Hygiene, includes items 10, 11, and 12
- Reach, includes items 13 and 14
- Grip, includes items 15, 16 and 17
- Activities, includes items 18, 19, and 20

The score for each category will be the single response within the category with the highest score (greatest difficulty). For example, in the "Eating" category, there are two answers (one for each item). If "Cut your food with a knife or fork" is marked as "3" and "Lift a full cup or glass to your mouth" is marked as "0", then the score for the "Eating" category would be "3" (the response indicating the greatest difficulty within the category).

If a component question is left blank or the response is too ambiguous to assign a score, then the score of that category will be determined by the remaining completed question(s).

However, if **any** "aids or devices" and/or "help from another person" items at the bottom of each page are checked, the category to which they apply will be adjusted upward to "2". If the basic score is **already** "2" or "3", the score remains unchanged. The checking of the items "Aids or devices" and "help from another person" can **only** change a category's score to "2"; they do **not** change the score to a "1" or a "3".

The score for the disability index will be the mean of the eight category scores. If more than two of the categories, or 25%, are missing, scale will not be scored. Otherwise, divide the sum of the categories by the number of answered categories. The higher score indicates greater disability.

Errors occurred during the translation to some languages of the questionnaire in one of the items of the 4-point scale. This could have led to a misinterpretation of the remaining answer options, affecting patients responses to 20 questions. Therefore the answers of the affected languages and software version are flagged. These flagged answers will not be used in any analyses and only be listed with a flag and corresponding footnote.

**HAQ-DI response** is defined by an improvement of at least 0.3 score points compared to baseline.

### **Major clinical response**

Major clinical response is defined as continuous six-months of ACR70 response for a subject.

### **DAS28, low disease activity and remission**

The Disease Activity Score (DAS) is a combined index to measure the disease activity in patients with RA. It has been extensively validated for its use in clinical trials in combination with the EULAR response criteria.

The DAS28 is a measure of disease activity based on Swollen and Tender Joint Counts, CRP or ESR, and the Patient Global Assessment of PsA disease activity.

The following 28 joints will be assessed for tenderness and swelling: metacarpophalangeal IV (10), thumb interphalangeal (2), hand proximal interphalangeal II-V (8), wrist (2), elbow (2), shoulders (2), and knees (2).

The following formulas can be used to calculate the DAS28 with CRP (mg/L) or ESR (mm/hour).

$$\text{DAS28-CRP} = 0.56 * \text{sqrt}(\text{TJC28}) + 0.28 * \text{sqrt}(\text{SJC28}) + 0.36 * \ln(\text{CRP}+1) + 0.014 * \text{PGA} + 0.96$$

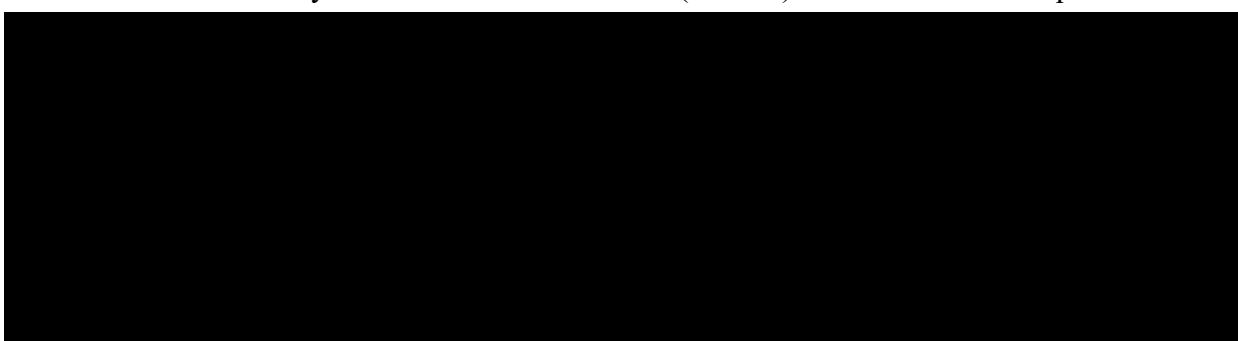
$$\text{DAS28-ESR} = 0.56 * \text{sqrt}(\text{TJC28}) + 0.28 * \text{sqrt}(\text{SJC28}) + 0.70 * \ln(\text{ESR}) + 0.014 * \text{PGA}$$

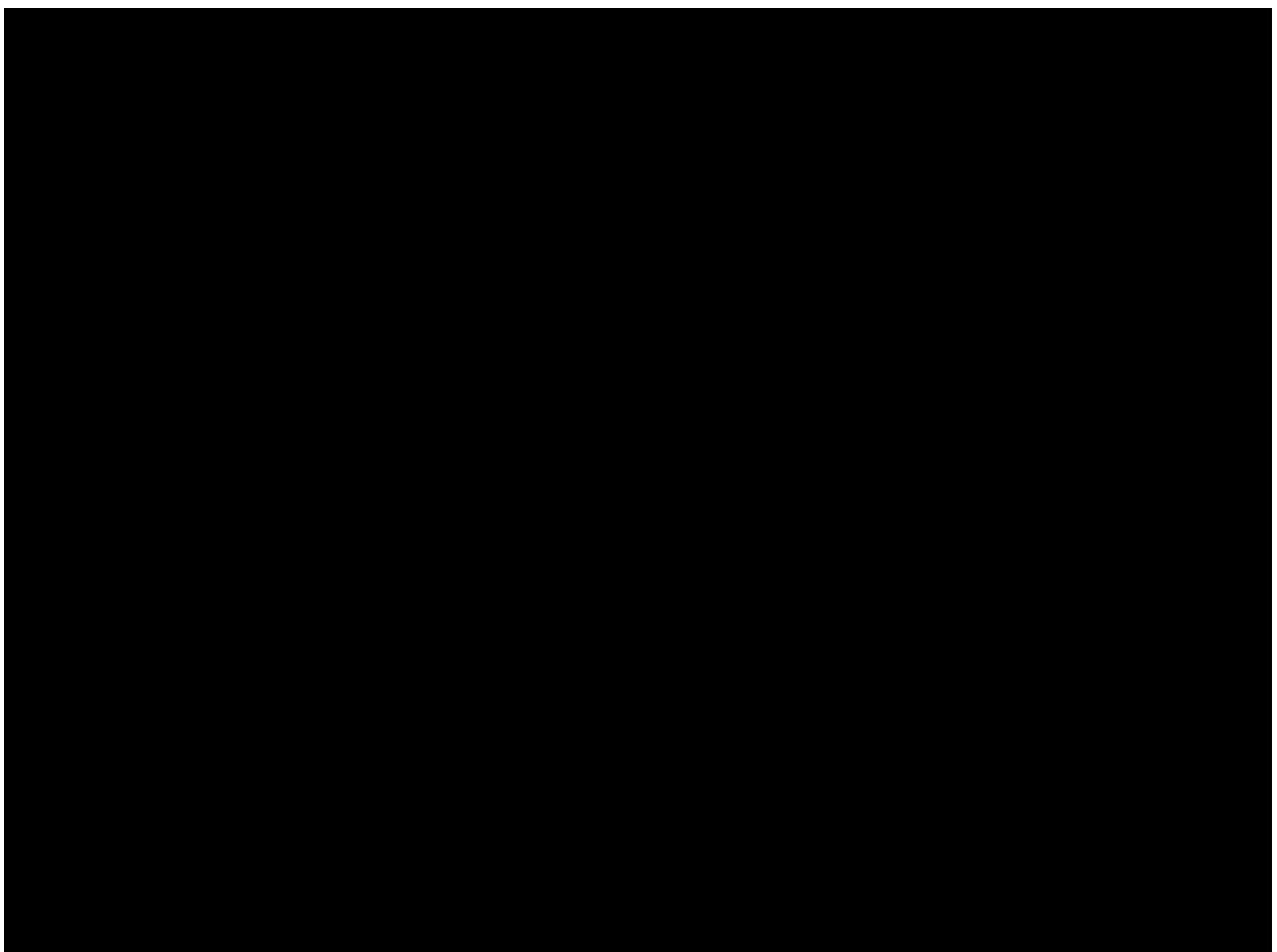
*TJC28: 28 Tender joint count; SJC28: 28 Swollen joint count; CRP: C-reactive protein; PGA: Patient Global Assessment*

If any component measurement is missing, DAS28 will be missing.

DAS28-CRP will be primarily used for the interpretation of the outcome; DAS-ESR is considered supportive.

DAS28-CRP (or ESR) remission is defined as a DAS28-CRP (or ESR) index score less than 2.6. Low disease activity is defined as DAS28-CRP (or ESR) index less than or equal to 3.2.





## **ACR Components**

### ***Tender 78 joint count and swollen 76 joint count***

The 78 joints assessed for tenderness include the 2 temporomandibular, 2 sternoclavicular, 2 acromioclavicular joints, 2 shoulders, 2 elbows, 2 wrists, 2 first carpometacarpal, 10 metacarpophalangeal, 10 proximal interphalangeal, 8 distal interphalangeal joints of the hands, the 2 hip, 2 knee, 2 talo-tibial, 2 mid-tarsal, 10 metatarsophalangeal, 10 proximal interphalangeal, and 8 distal interphalangeal joints of the feet. All of these except for the hips are assessed for swelling. Joint tenderness and swelling are to be graded present (1) or absent (0). Synovial fluid and/or soft tissue swelling but not bony overgrowth represents a positive result for swollen joint count. Dactylitis of a digit in the foot or hand counts as one tender and swollen joint. These aspects are to be implemented during data capture and will not be addressed during programming.

If the number of joints for which data were available (e.g., T) is less than 78/76 for the tender/swollen joint assessment, the number of tender/swollen joints (e.g., t) will be scaled up proportionately (i.e.,  $78*t/T$  or  $76*t/T$  for tender or swollen joint count).

### ***Patient's assessment of PsA Pain***

The patient's assessment of pain will be performed using 100 mm visual analog scale (VAS) ranging from "no pain" to "unbearable pain" after the question "*Please indicate with a vertical mark (|) through the horizontal line the most pain you had from your psoriatic arthritis today*".

### **Patient's global assessment of disease activity**

The patient's global assessment of disease activity will be performed using 100 mm VAS ranging from "very good" to "very poor", after the question "*Considering all the ways psoriatic arthritis affects you, please indicate with a vertical mark (|) through the horizontal line how well you are doing today*".

### **Physician's global assessment of PsA disease activity**

The physician's global assessment of disease activity will be performed using 100 mm VAS ranging from no disease activity to maximal disease activity, after the question "*Considering all the ways the disease affects your patient, draw a line on the scale for how well his or her condition is today*". To enhance objectivity, the physician must not be aware of the specific patient's global assessment of disease activity, when performing his own assessment on that patient.<sup>4</sup>

### **Erythrocyte sedimentation rate (ESR)**

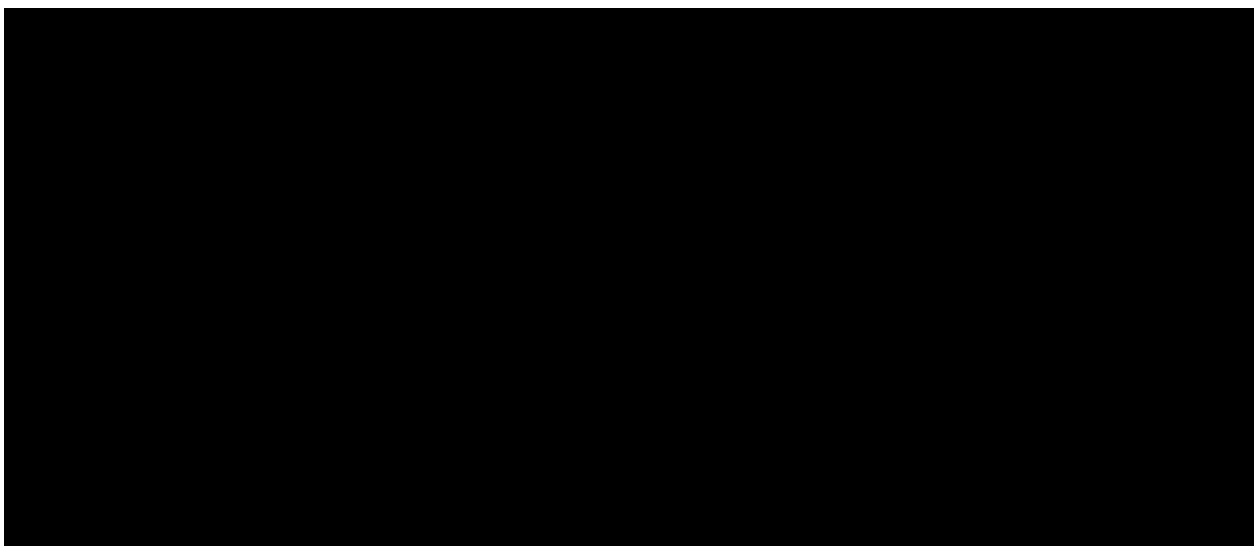
Blood for ESR, which is helpful in diagnosing inflammatory diseases and is used to monitor disease activity and response to therapy, will be obtained at scheduled visits.

### **High-sensitivity C-reactive protein (hsCRP)**

Blood for this assessment will be obtained in order to identify the presence of inflammation, to determine its severity, and to monitor response to treatment. Since the results of this test may unblind study personnel, results from the central lab will be provided for screening and baseline only. The hsCRP results from samples collected during the treatment period will be revealed following database lock only.

### **Radiographic assessments**

See Section 2.12.4.1 on Joint/bone structural damage for a description of the derivation of the vdH-mTSS.



### **Presence of enthesitis**

If enthesitis is present with any of the 6 sites (lateral epicondyle humerus L + R, proximal achilles L + R and medial condyle femur L + R), the patient is counted as a patient with enthesitis.

### **Dactylitis count**

The dactylitis count is the number of fingers and toes with dactylitis based on the LDI, with a range of 0-20.

### **Presence of dactylitis**

If dactylitis according to the LDI is present with any finger or toe, the patient is counted as a patient with dactylitis.

### **Psoriasis Area and Severity Index (PASI)**

The PASI assessment will be conducted for subjects in whom at least 3% of the body surface area (BSA) was affected by psoriatic skin involvement at baseline (Psoriasis Subset). The PASI assesses the extent of psoriasis on four body surface areas (head, trunk and upper and lower limbs) and the degree of plaque erythema, scaling and thickness. A PASI score will be derived as indicated in [Table 2-7](#). The head, trunk, upper limbs and lower limbs are assessed separately for erythema, thickening (plaque elevation, induration), and scaling (desquamation). The average degree of severity of each sign in each of the four body regions is assigned a score of

0-4. The area covered by lesions on each body region is estimated as a percentage of the total area of that particular body region. Further practical details help the assessment:

1. The neck is assessed as part of the head.
2. The axillae and groin are assessed as part of the trunk.
3. The buttocks are assessed as part of the lower limbs.
4. When scoring the severity of erythema, scales should not be removed.

**Table 2-7 The PASI scoring system**

Body region	Erythema (E)	Thickening (plaque elevation, induration, I)	Scaling (desquamation) (D)	Area score (based on true area %, A)*
Head (H)**	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0 = no involvement 1 = >0-< 10% 2 = 10-<30% 3 = 30-<50% 4 = 50-<70% 5 = 70-<90% 6 = 90-100%
Trunk, (T)***	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0 = no involvement 1 = >0-< 10% 2 = 10-<30% 3 = 30-<50% 4 = 50-<70% 5 = 70-<90% 6 = 90-100%
Upper limbs (U)	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0 = no involvement 1 = >0-< 10% 2 = 10-<30% 3 = 30-<50% 4 = 50-<70% 5 = 70-<90% 6 = 90-100%
Lower limbs (L)****	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0=none 1=slight 2=moderate 3=severe 4=very severe	0 = no involvement 1 = >0-< 10% 2 = 10-<30% 3 = 30-<50% 4 = 50-<70% 5 = 70-<90% 6 = 90-100%

Percentage (not score) of body region (not whole body) affected will be entered in the eCRF

\*\*Neck is assessed as part of the Head (H) body region.

\*\*\*Axillae and groin are assessed as part of the Trunk (T) body region.

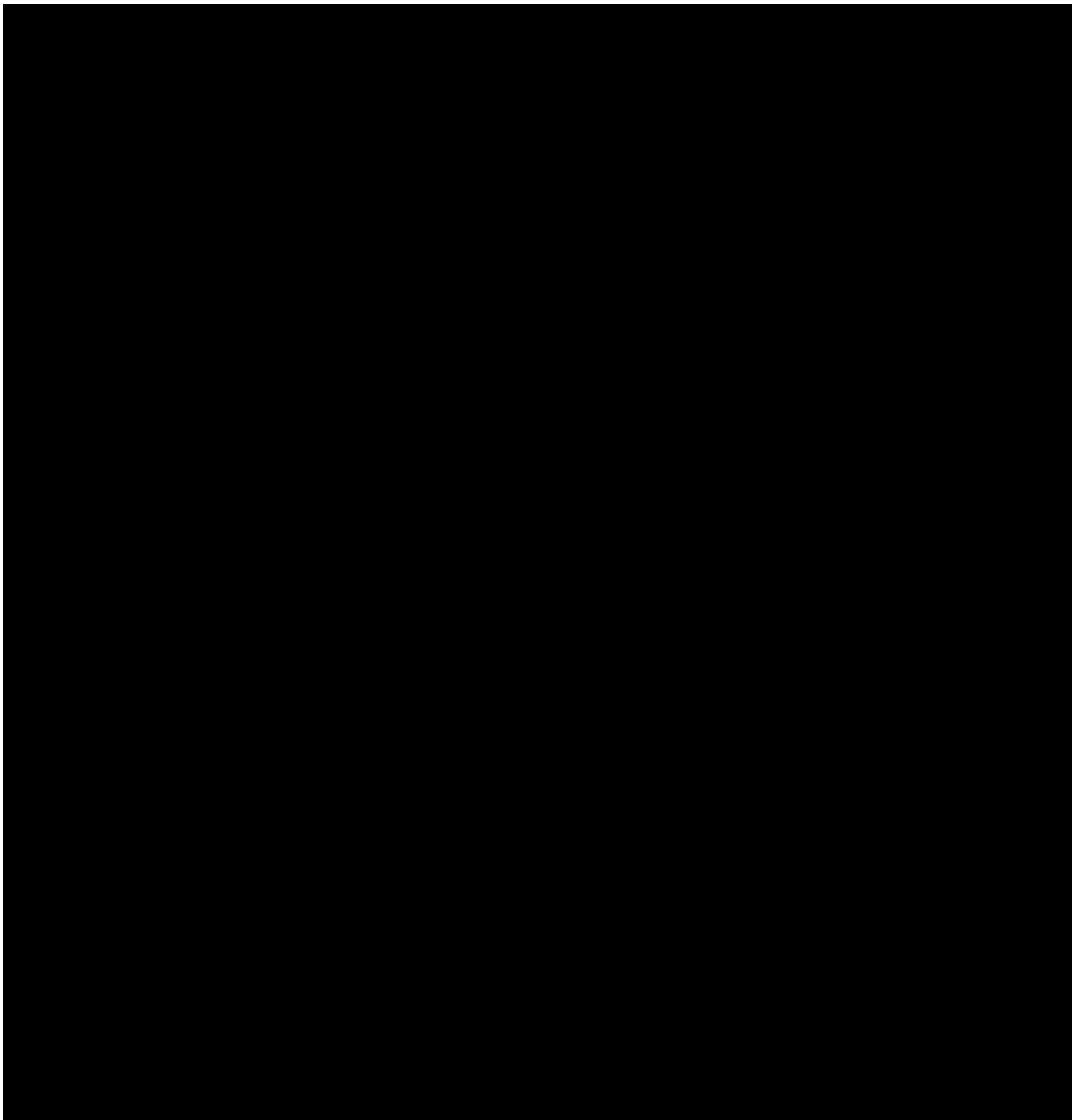
\*\*\*\*Buttocks are assessed as part of the Lower limbs (L) body region.

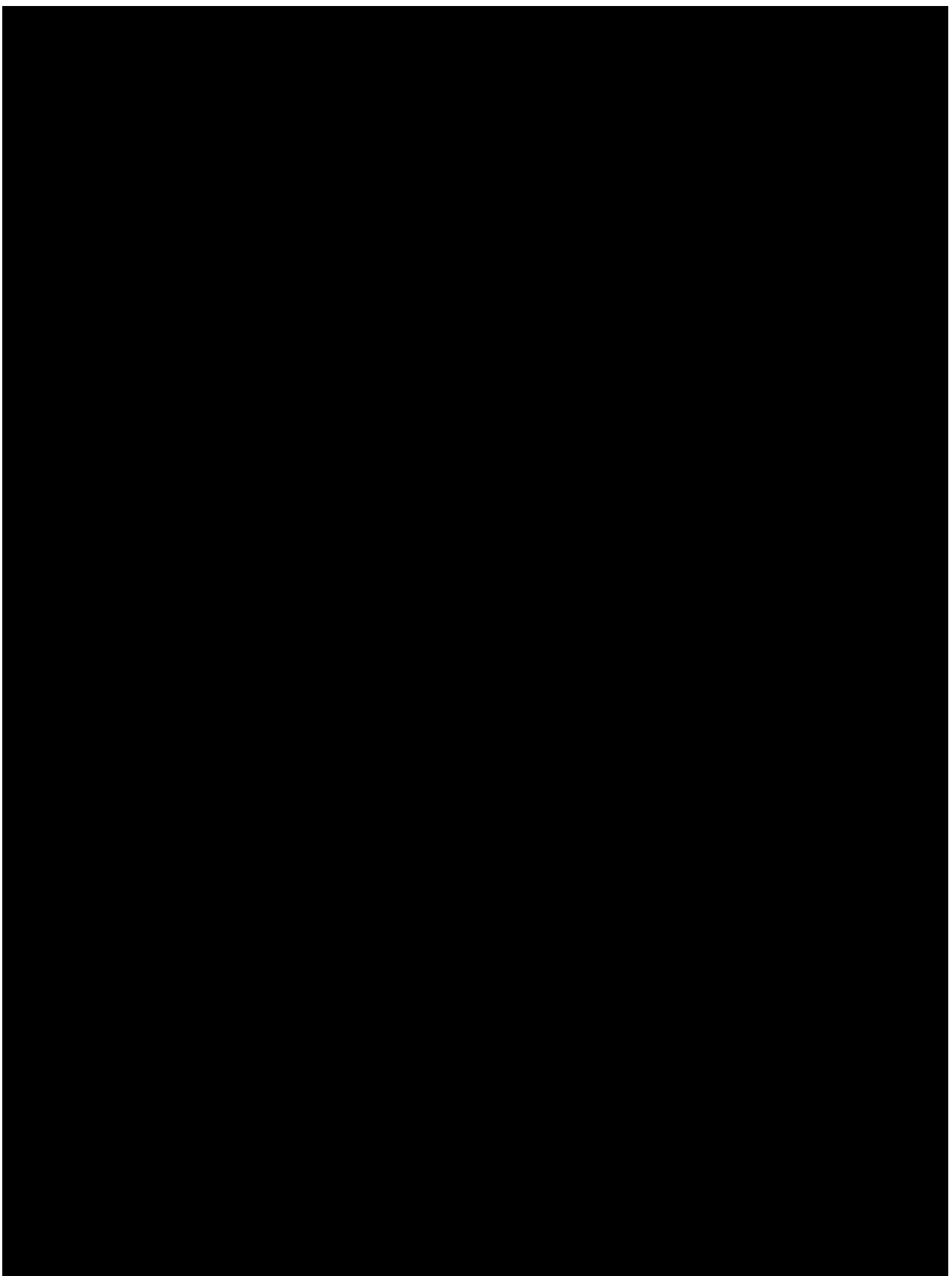
Because the head and neck, upper limbs, trunk and lower limbs correspond to approximately 10%, 20%, 30% and 40% of the body surface area, respectively, the PASI score is calculated using the formula:

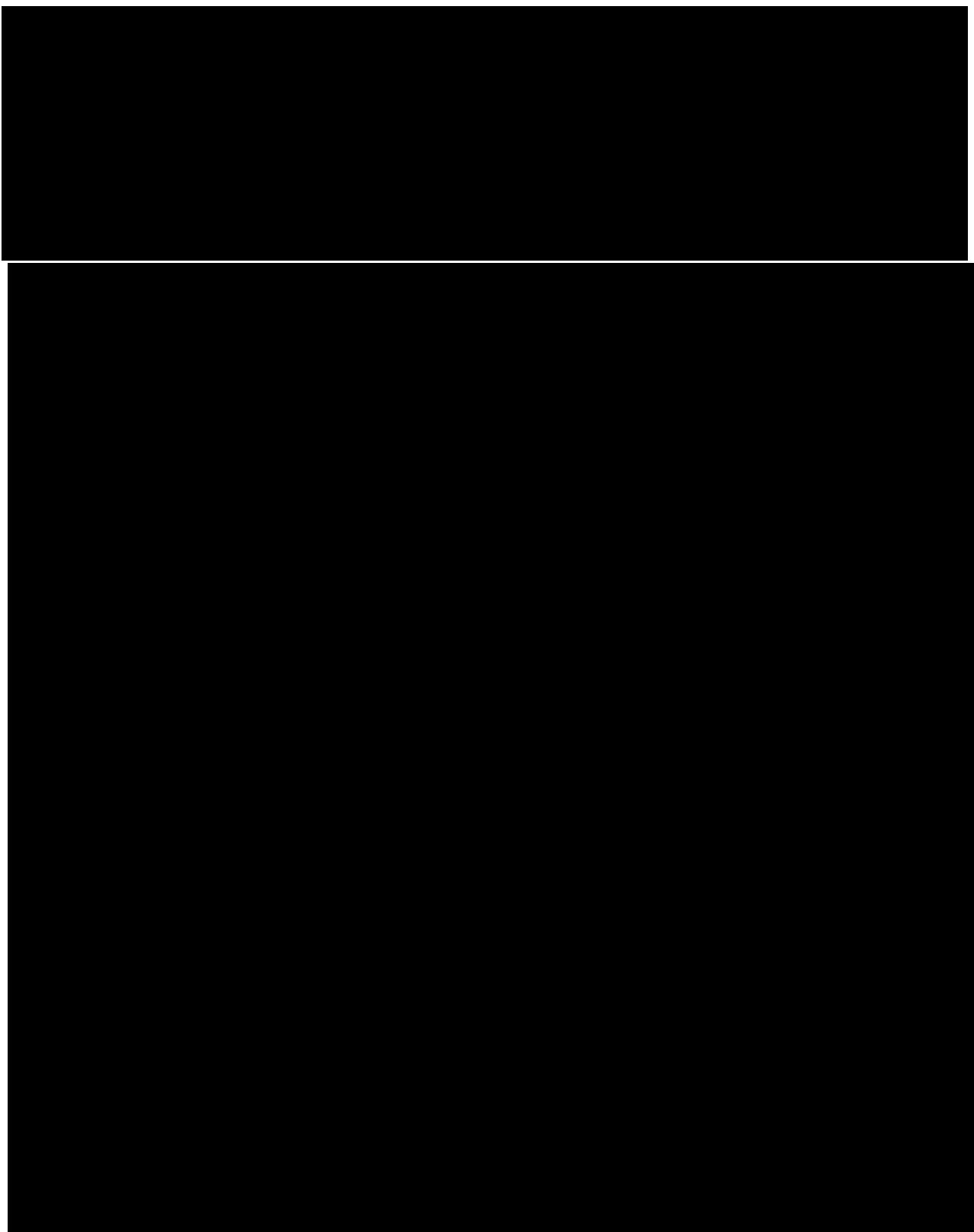
$$\text{PASI} = 0.1(E_H + I_H + D_H)A_H + 0.2(E_U + I_U + D_U)A_U + 0.3(E_T + I_T + D_T)A_T + 0.4(E_L + I_L + D_L)A_L$$

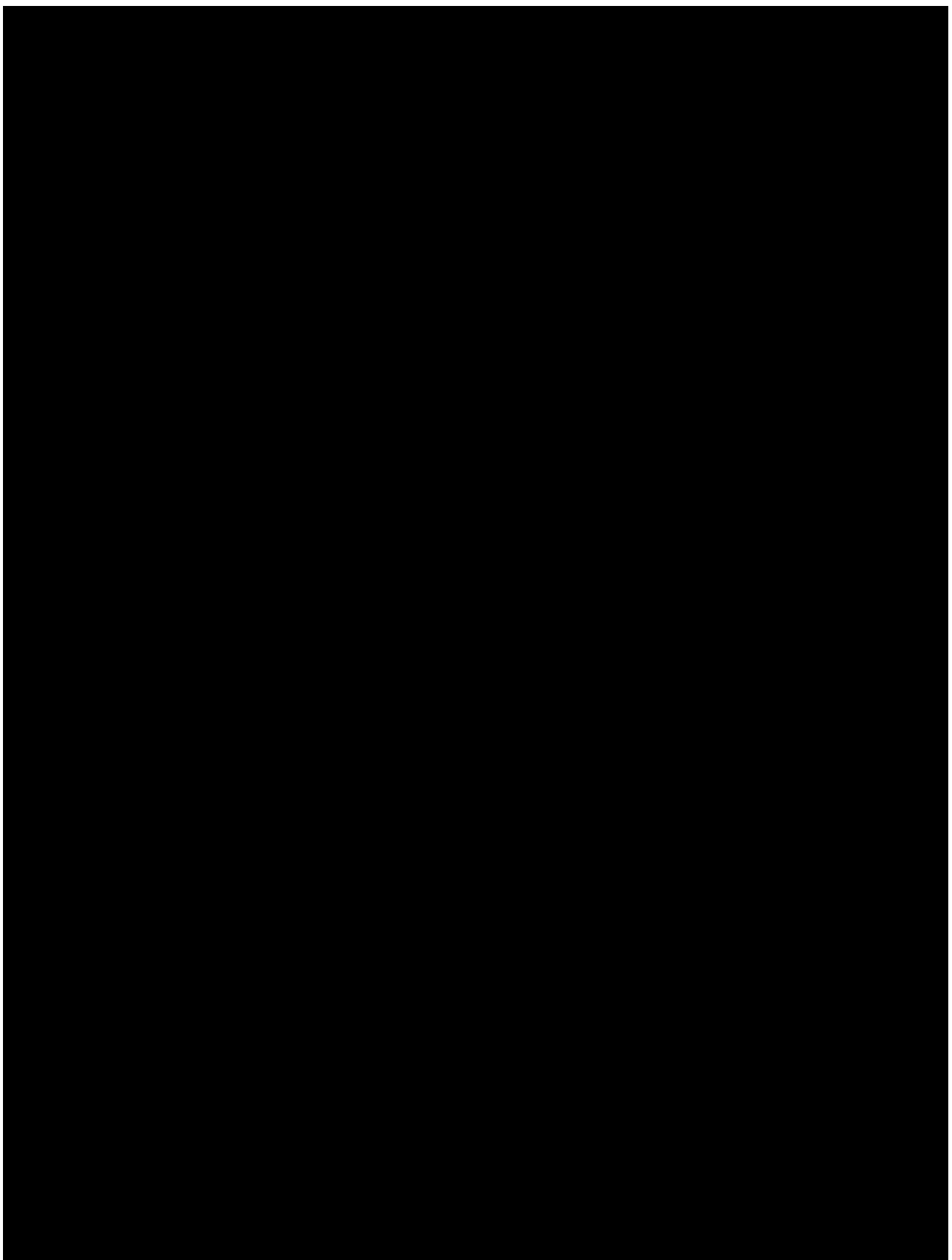
The keys for the letters are provided in [Table 2-7](#).

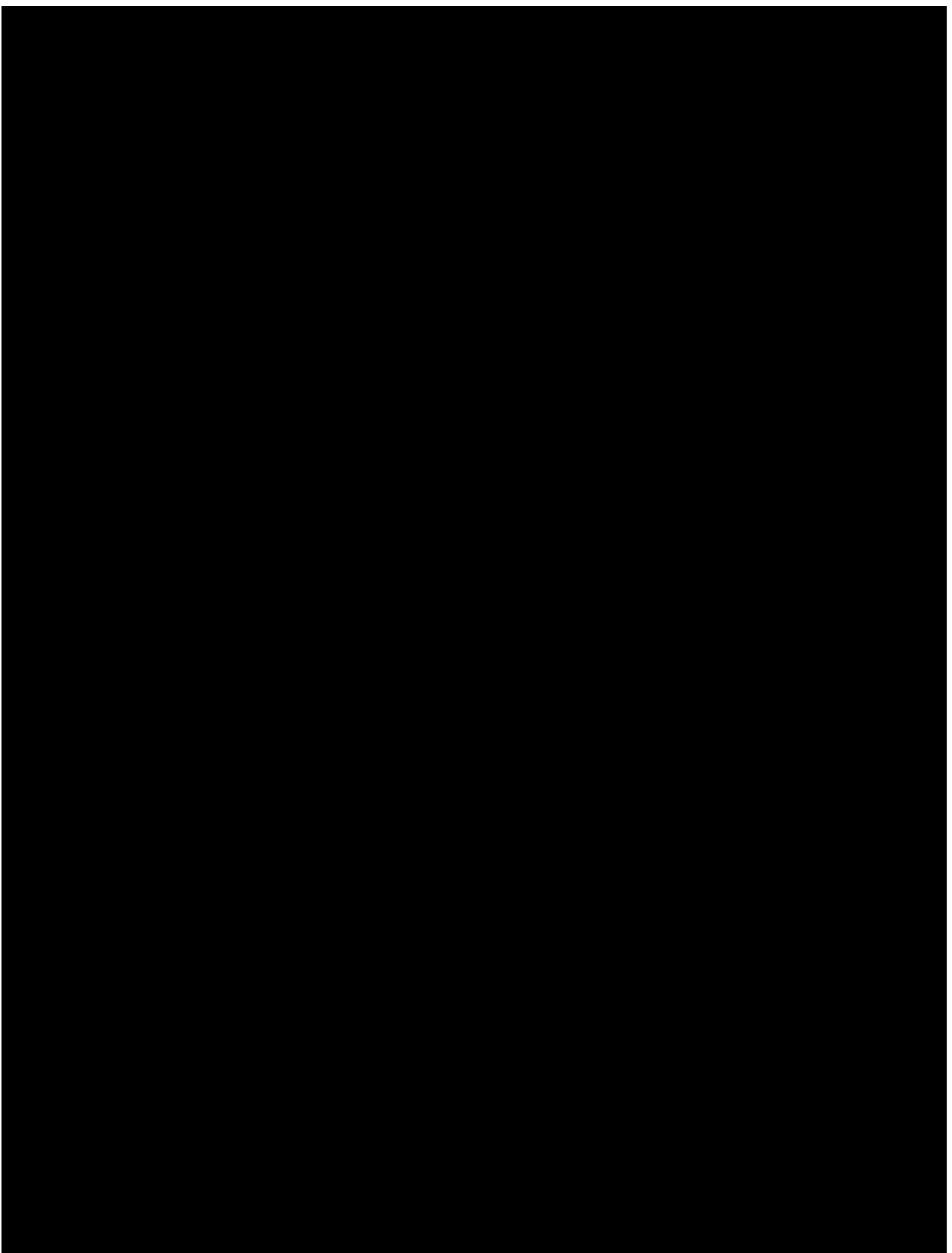
PASI scores can range from a lower value of 0, corresponding to no signs of psoriasis, up to a theoretic maximum of 72.0. The total score comes from eCRF.











## **2.13 Interim analysis**

The primary endpoint analysis will be performed after all subjects have completed the Week 24 visit. The investigators, site personnel and monitors will continue to remain blinded to the treatment each subject received at randomization until database lock for Week 52 analysis.

Subsequent to the primary endpoint analysis, additional analyses are planned for regulatory submission and/or publication purposes after subjects have completed the Week 52 assessments. The final analysis will be conducted after all subjects complete the study. Additional analyses may be performed to support interactions with health authorities, as necessary.

## **3 Sample size calculation**

A 5% two-sided type I error rate will be used to control for type I error. Three secukinumab regimens will be tested versus placebo with respect to the primary endpoint (ACR20 response at Week 24), thus the type-I-error will be split to 1.67% two-sided for each comparison. Sample sizes will be based on this type I error assumption. With an assumed placebo rate of 15% and secukinumab 51% (FUTURE 2), the study is over 99% powered to detect a treatment difference of ACR20 in the FAS population, assuming 220 subjects in a secukinumab treatment group, and 330 subjects in the placebo group (Fisher's exact test, nQuery 7.0). This power is applicable for all three dose regimens, as the sample size is driven by structural endpoint.

Power for secondary variables was calculated using a two-sided 1.67% type I error. With an assumed placebo rate of 7% and secukinumab 35% (FUTURE 2), the study is over 99% powered to detect a treatment difference of ACR50 in the full FAS population, assuming 220 subjects in a secukinumab treatment group, and 330 subjects in the placebo group (Fisher's exact test, nQuery 7.0).

For structural endpoint, historical data (FUTURE 1) showed a standard deviation of 1.132 on active treatment and 2.435 on placebo at Week 24, and a difference of 0.52. Using the above assumptions, there is 83% power to show statistically significant differences assuming 220 subjects in a secukinumab group and 330 subjects in the placebo group (Satterthwaite t-test, nQuery 7.0).

A standard deviation of approximately 0.49 and a treatment difference of 0.17 has been observed for the change from baseline at Week 24 in HAQ-DI<sup>®</sup> (FUTURE 2). Using those assumptions, the study has approximately 94% power to detect a difference between secukinumab and placebo (Two group t-test, nQuery Advisor 7.0), assuming 220 subjects in a secukinumab treatment group and 330 subjects in the placebo group.

For the presence of dactylitis at Week 24 in the subset of subject who have dactylitis at BSL, with an assumed placebo rate of 85% and secukinumab 50% (FUTURE 2), there is about 99% power to show statistically significant difference between secukinumab (110) and placebo (165 subjects), assuming 50% subject have dactylitis at BSL (Fisher's exact test, nQuery 7.0).

For the presence of enthesitis at Week 24 in the subset of subject who have enthesitis at BSL, with an assumed placebo rate of 79% and secukinumab 58% (FUTURE 2), there is about 94% power to show statistically significant difference between secukinumab (132 subjects) and

placebo (198 subjects), assuming 60% subjects have enthesitis at BSL (Fisher's exact test, nQuery 7.0).

With an assumed placebo rate of 16% and secukinumab 48% (FUTURE 2), the study is over 99% powered to detect a treatment difference of PASI75 in the FAS population, assuming 110 subjects in a secukinumab treatment group, and 165 subjects in the placebo group. It is assumed that about 50% of enrolled subjects have  $\geq 3\%$  skin involvement.

A difference of 0.62 and standard deviation of 1.04 has been observed for the change from baseline in DAS28-CRP (FUTURE 2). With these assumptions, the study has over 99% power to detect a difference between secukinumab and placebo (Two group t-test, nQuery Advisor 7.0), assuming 220 subjects in a secukinumab treatment group, and 330 subjects in the placebo group.

A standard deviation of approximately 7.19 and a treatment difference of 4.44 has been observed for the change from baseline at Week 24 in SF36-PCS (FUTURE 2). Using those assumptions, the study has over 99% power to detect a difference between secukinumab and placebo (Two group t-test, nQuery Advisor 7.0), assuming 220 subjects in a secukinumab treatment group, and 330 subjects in the placebo group.

With an assumed placebo rate of 9% and secukinumab 33% (FUTURE 2), the study is over 99% powered to detect a treatment difference of PASI90 in the FAS population, assuming 110 subjects in a secukinumab treatment group, and 165 subjects in the placebo group. It is assumed that about 50% of enrolled subjects have  $\geq 3\%$  skin involvement.

## **4 Appendix**

### **4.1 Visit Windows and Cut-off Dates**

#### **4.1.1 Visit windows**

When visit windows are used, all visits will be re-aligned, i.e., they will be mapped into one of the visit windows. E.g., if the *Week 4* visit of a subject is delayed and occurs on Day 46 instead of on Day 29, say, it will be re-aligned to visit window *Week 8*. In the case of major deviations from the visit schedule, or due to unscheduled visits, several assessments of a subject may fall in a particular visit window (either scheduled or unscheduled). Statistical approaches to handle multiple assessments in a given visit window are specified in [Table 4-2](#).

For lab/ECG/vital signs, follow-up (F/U) visit is excluded from analysis visit mapping window. Only assessments that come as F/U nominal visit will be directly assigned as analysis F/U visit. Other assessments that are beyond the last on-treatment visit window (W104) or after nominal F/U visit date won't be mapped to any analysis visit. F/U visit will not be included in the summary tables by visit.

Of note, subjects are allowed to have gaps in visits. All data collected will be displayed in listings.

**Table 4-1 Analysis visit windows**



The following rules are used to determine the window for an applicable visit post baseline: “Lower limit” = “upper limit of prior applicable visit” + 1. “Upper limit” = “target day of current visit” + integer part of (“target day of next applicable visit” – “target day of current visit”)/2. Lower limit of the first applicable visit is always Day 2. Day 1 is the date of the first dose of randomized study treatment

The mapping described above applies to all visits (not just scheduled visits). Repeat and/or unscheduled visits (which will be numbered in the database according to new NCDS standards) will be mapped for analysis purposes in the same way as scheduled visits. This leaves the possibility, then, for multiple measurements within an analysis window. The following conventions will be used to determine the appropriate measurement to be summarized in the event of multiple measurements within a visit window.

**Table 4-2 Rules for flagging variables**

Timing of measurement	Type of data	Rule
Baseline	All data	<p>The last non-missing measurement made prior to or on the date of administration of the first dose of study treatment (the reference start date / Day 1). If a patient did not receive any dose of study treatment then the randomization date will be used. Only date part is considered if just one assessment on Day 1.</p> <p>If there are multiple assessments on Day 1, following rules will apply:</p> <ul style="list-style-type: none"><li>a. If assessment time exists,<ul style="list-style-type: none"><li>• select the last available measurement prior to reference start date/time considering time;</li><li>• if no measurement prior to reference start date/time considering time, select the earliest measurement post reference start date/time considering time .</li></ul></li><li>b. If assessment time does not exist, select the available measurement from the lowest CRF visit number.</li></ul> <p>For X-ray, a baseline value is the last measurement prior to dosing if available. Otherwise, take the first value within 30 days post dosing.</p> <p>For MRI, a baseline value is the last measurement prior to dosing if available. Otherwise, take the first value within 7 days post dosing.</p>
Post-baseline efficacy	All data	<p>For visits without switch of treatment in the window, the measurement closest to the target will be used. In the event two measurements are taken equally apart (e.g. 1 before target date and 1 after) the first one will be used.</p> <p>For visits during which the patient switches from placebo to AIN the following will be done based on whether or not the patient met the rescue criteria:</p> <ul style="list-style-type: none"><li>• If the analysis visit window is &lt;= week 16(for non-responders) or week 24 (for responders), then:</li></ul>

Timing of measurement	Type of data	Rule
		<ul style="list-style-type: none"><li>• If available, the closest measurement to the target date which is ON or BEFORE the switch date will be used (i.e. the closest measurement to target which is on placebo).</li><li>• If there are no data on or before the switch then the closest measurement to the target date after the switch will be used.</li><li>• If the analysis visit window is &gt; week 16(for non-responders) or week 24 (for responders), then<ul style="list-style-type: none"><li>• If available, the closest measurement to the target date which is AFTER the switch date will be used (i.e. the closest measurement to target which is on AIN).</li><li>• If there are no data AFTER the switch then the closest measurement to the target date before the switch will be used.</li></ul></li></ul> <p>Cases where the same parameter is recorded more than once on the same date will be handled as follows:</p> <ul style="list-style-type: none"><li>• If time of completion exists the earliest measurement will be used;</li><li>• If time does not exist the measurement from the lowest CRF visit number will be used.</li></ul>
Post-baseline safety	Summary visit information (e.g. lab, ECG, etc.)	<p>For visits without switch of treatment in the window, the measurement closest to the target will be used. In the event two measurements are taken equally apart (e.g. 1 before target date and 1 after) the first one will be used.</p> <p>For visits during which the patient switches from placebo to AIN the following will be done based on whether or not the patient met the rescue criteria:</p> <ul style="list-style-type: none"><li>• If the analysis visit window is &lt;= week 16(for non-responders) or week 24 (for responders), then:<ul style="list-style-type: none"><li>• If available, the closest measurement to the target date which is ON or BEFORE the switch date will be used (i.e. the closest measurement to target which is on placebo).</li><li>• If there are no data on or before the switch then the closest measurement to the target date after the switch will be used.</li></ul></li><li>• If the analysis visit window is &gt; week 16(for non-responders) or week 24 (for responders), then<ul style="list-style-type: none"><li>• If available, the closest measurement to the target date which is AFTER the switch date will be used (i.e. the closest measurement to target which is on AIN).</li><li>• If there are no data AFTER the switch then the closest measurement to the target date before the switch will be used.</li></ul></li></ul>

Timing of measurement	Type of data	Rule
		Cases where the same parameter is recorded more than once on the same date will be handled as follows: <ul style="list-style-type: none"><li>• If time of completion exists the earliest measurement will be used;</li><li>• If time does not exist the measurement from the lowest CRF visit number will be used.</li></ul>
Post-baseline safety	Notable abnormalities (e.g. lab)	The most extreme measurement in the window will be used. Note this means a patient can have a notably high and notably low measurement within a window.

#### 4.1.2 Cut-off dates

The cut-off dates for various DBL analyses will be defined as follows:

Week 24 DBL and analysis: For all domains with visit numbers (e.g. LV, VS, EG etc.), Week 24 visit number will be used for cut-off. For event domains without visit numbers (e.g. AE, CM, MH etc.), the date in DS (the disposition date to complete or discontinue Week 24) will be used for cut-off.

Week 52 DBL and analysis: For all domains with visit numbers (e.g. LV, VS, EG etc.), Week 52 visit number will be used for cut-off. For event domains without visit numbers (e.g. AE, CM, MH etc.), the date in DS (the disposition date to complete or discontinue Week 52) will be used for cut-off.

Week 112 / End of Study DBL and analysis: All data up to Day 799 will be included in the end of study analysis..

### 4.2 Detailed on implementation of statistical methodology and assumptions

#### 4.2.1 Analysis of continuous data

##### 4.2.1.1 Summary statistics for continuous data

Summary statistics (including N, mean, standard deviation, minimum, lower quartile, median, upper quartile, maximum) will be provided for continuous data by visit and treatment group.

##### 4.2.1.2 Mixed-effects repeated measures model

Endpoints with continuous data type expected to be normally distributed (e.g, DAS28) will be analyzed using a mixed-effects repeated measures model (MMRM) with treatment, stratification factor and analysis visit as factors; and weight, baseline value, treatment by visit and baseline by visit interactions as covariates. An unstructured covariance structure will be assumed for this model. The significance of the treatment effects for secukinumab regimens at different analysis visits will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at the appropriate analysis visits.

SAS code for mixed model:

```
proc mixed data=aaa;
class TRT USUBJID AVISITN STRATA;
model CHG=TRT STRATA AVISITN WEIGHT BASE TRT*AVISITN BASE*AVISITN
/ s ddfm=kr;
lsmeans TRT*AVISITN / diff cl;
repeated AVISITN / type=un subject=USUBJID;
Run;
```

In case the MMRM model does not converge the following sequential steps will be used:

1. change ddfm=kr to ddfm=bw. If still no convergence, perform step 2.
2. change type=un to type=cs. If still no convergence, perform step 3.
3. remove covariates in the following order until convergence: WEIGHT, BASE\*AVISITN, STRATA.

#### **4.2.1.3 Non-parametric analysis of covariance**

A non-parametric ANCOVA model (Koch 1998) will be used for the endpoints that are not normally distributed, e.g. X-ray, MRI, et. ctrl. The macro NParCov3 will be used, see Zink and Koch 2012.

For continuous response variable, the macro call will be as follows:

```
%NParCov3(OUTCOMES = response, COVARS = weight baseline, C=1,
HYPOTH = ALT, STRATA = TNF $\alpha$  status, TRTGRPS = treatment, TRANSFORM = NONE,
COMBINE = FIRST, DSNIN = RESP, DSNOUT = OUTDAT);
```

Data set \_OUTDAT\_DEPTEST provides results for the treatment difference, and \_OUTDAT\_CI provides a 95% confidence interval for the treatment estimate.

For binary response variable, the macro call will be as follows:

```
%NParCov3(OUTCOMES = response, COVARS = weight, C = 1, HYPOTH = ALT, STRATA
= TNF $\alpha$  status, TRTGRPS = treatment, TRANSFORM = LOGISTIC,
COMBINE = FIRST, DSNIN = RESP, DSNOUT = OUTDAT);
```

The odds ratio and confidence interval are to be obtained from \_OUTDAT\_RATIOCI.

#### **4.2.2 Analysis of binary (and categorical) data**

##### **4.2.2.1 Analyses of Frequencies**

Summary statistics for discrete variables will be presented in contingency tables and will include count and frequency in each category. For n subjects, each at risk to experience a certain event with probability  $\pi$ , the crude incidence is estimated as  $p=x/n$ , where x is the number of subjects with the event.

If applicable, 95% confidence intervals for the relative frequencies will be derived as well based on the score method including continuity correction [Newcombe (1998)]:

With  $z$  as  $(1-\alpha/2)$ -quantile of the standard normal distribution (SAS:  $z=\text{PROBIT}(1-\alpha/2)$ ),  $n$  as total number of subjects (i.e. number of subjects in the denominator), and  $p$  as estimated crude incidence (number of subjects with event /  $n$ ) it is  $q=1-p$

Then the lower limit is

$$L = \max\left(0, \frac{2np + z^2 - 1 - z\sqrt{z^2 + 2 - \frac{1}{n} + 4p(nq+1)}}{2(n+z^2)}\right)$$

and the upper limit is

$$U = \min\left(1, \frac{2np + z^2 + 1 + z\sqrt{z^2 + 2 - \frac{1}{n} + 4p(nq-1)}}{2(n+z^2)}\right).$$

In addition, if  $L > p$  then  $L = p$  and if  $U < p$  then  $U = p$ .

For binary response variables (e.g. for ACR20/50/70, HAQ-DI responder, PASI 75, IGA response) the placebo-adjusted response rates including 95% confidence interval will be derived.

SAS code for risk difference:

```
Proc freq data=acr order=formatted;
```

```
Tables response*trt/ riskdiff;
```

```
Run;
```

(Note the response value should be sorted with '1' ahead of '0'.)

Fisher's exact test will be applied to rare events (e.g., MCR), pairwise treatment group comparisons to placebo or active controls.

SAS code for Fisher's exact test:

```
Proc freq data=mcr order=formatted;
```

```
Tables response*trt/Fisher;
```

```
Run;
```

If appropriate, an exact  $100*(1-\alpha)\%$  confidence interval (Clopper-Pearson 1934) will be obtained by using the SAS procedure PROC FREQ with the EXACT BINOMIAL statement. However, the confidence interval derived via the score method including continuity correction will be the default in safety analyses.

Figures will be provided for primary and secondary variables, with means and 95% confidence intervals displayed across time for all the treatment groups.

#### 4.2.2.2 Logistic regression

Certain binary outcome variables, e.g. response outcomes, will be evaluated using a logistic regression model with treatment regimen, weight, stratum if applicable. Odds ratios will be computed for comparisons of AIN457 regimens versus control(s) utilizing the logistic regression model fitted.

SAS code for logistic regression:

```
Proc logistic data=aaa;  
  Class TRT STRATA / param=glm;  
  Model AVAL = TRT WEIGHT STRATA;  
  Lsmeans TRT / diff cl exp;  
  Ods output diffs=lsm_diff;  
  Run;
```

Logistic regression will be applied to response variables at each visit.

In cases where logistic regression doesn't converge, Fisher's exact test will be applied for comparisons of between AIN457 doses. In this case, no odds ratios or confidence intervals will be estimated, but p-values may be calculated.

```
Proc freq data=aaa;  
  Table TRT * AVAL / fisher;  
  Where TRT in ("AIN457 xx mg" "Placebo");  
  Run;
```

#### 4.2.2.3 Cochran-Mantel-Haenszel Test

The CMH test will be performed using the SAS procedure PROC FREQ with the CMH option.

#### 4.2.3 Multiple Imputation

A multiple imputation will be performed based on MAR by treatment group for baseline weight, baseline and post-baseline of each parameter for visits up to the primary time point (e.g Week 16 or 24) using Markov Chain Monte Carlo (MCMC) method with EM algorithm.

Impute the missing values 100 times (NIMPUTE) with a seed=457<studycode> as shown below:

```
proc mi data= out=imp minmaxiter=10000000 nimpute=100 seed=4572342;  
  by trt;  
  var weight_base strata var1_base var1_week1-var1_week24;  
  mcmc chain=multiple initial=em;  
  run;
```

If needed repeat for each component necessary to calculate the final score, e.g. as follows:

```
proc mi data=imp out=imp2 minmaxiter=10000000 nimpute=100 seed=4572342;  
  by trt _imputation_ ;  
  var weight_base strata var2_base var2_week1-var2_week24;  
  mcmc chain=multiple initial=em;  
run;
```

Post-processing of out of range imputed values will be processed with the min and max values of each variable.

The score and ACR response can now be calculated based on the complete data. The response rate will be calculated for each imputation and then combined using Rubin's rules.

In order to calculate the response rate for each imputation, PROC FREQ will be used as follows. Calculate binomial proportion and standard error for each imputation.

```
proc freq data=<ACR20>;  
  by treat visit _imputation_ ;  
  tables <response> / binomial (level=2 cl=wilson correct) ;  
  ods output BinomialProp=imp_bpr;  
run;
```

Transpose the dataset for subsequent use with PROC MIANALYZE.

```
proc transpose data=imp_bpr out=imp_trs(drop=_name_) ;  
  by treat visit _imputation_ ;  
  var nvalue1; id name1; idlabel labell;  
run;
```

Apply LOGIT transformation:  $y=\log(p/(1-p))$  and std. err. transformation: <new se> =  $se/(p*(1-p))$

```
data logit;  
  set imp_trs(rename=(_bin_=p e_bin=se));  
  by treat visit _imputation_ ;  
  lmean=log(p/(1-p));  
  lse=se/(p*(1-p));  
run;
```

The transformed binomial proportion estimates and standard errors are combined by applying Rubin's rules for multiple imputed data sets.

```
proc mianalyze data=logit;  
  by treat visit ;  
  modeleffects lmean;  
  stderr lse;  
  ods output ParameterEstimates=logitres;  
run;
```

The combined data should be transformed back using the following formula:  $p=1/(1+\exp(-y))$

```
data miexpress;  
  set logitres;  
  by treat visit ;  
  resti = 1/(1+exp(-estimate));  
  rlow = 1/(1+exp(-lclmean));  
  rupp = 1/(1+exp(-uclmean));  
run;
```

Of note, sometimes all responses may be imputed to 0 or 1 at a given combination of response variable, treatment group and visit. Such cases should be considered separately. The combined final response rate would be the same as the original response but the 95% CI will be undefined.

The odds ratio will be derived using GENMOD for each imputation, then combined using Rubin's rules again.

```
proc genmod data = acr20_mi descending;  
  by avisitn _imputation_;  
  class trt_ TNFRESN ;  
  model aval = trt_ TNFRESN weight / link=logit dist=bin;  
  lsmeans trt_ / diff;  
  estimate 'AIN457 150 mg vs No Load' trt_ 1 -1 0 0;  
  estimate 'AIN457 150 mg vs Placebo' trt_ 1 0 0 -1;  
  estimate 'AIN457 150 mg No Load vs Placebo' trt_ 0 1 0 -1;  
  estimate 'AIN457 300 mg vs Placebo' trt_ 0 0 1 -1;
```

```
ods output Estimates=imp_est;
run;
proc mianalyze data=imp_est;
  by avisitn trt_;
  modeleffects LBetaEstimate;
  stderr StdErr;
  ods output ParameterEstimates=_res;
run;
```

#### 4.2.4 Crude incidence and relativ risk estimates

##### 4.2.4.1 Odds ratio and 100\*(1- $\alpha$ )% confidence interval

For an investigational drug group with  $n_1$  subjects at risk, independent from the control group (e.g. placebo or comparator) with  $n_0$  subjects at risk, of whom  $x_1$  and  $x_0$  experience a certain event with probability  $\pi_1$  and  $\pi_0$  respectively, the odds ratio is estimated as

$$\frac{p_1/(1-p_1)}{p_0/(1-p_0)}$$
 with  $p_1 = x_1/n_1$  and  $p_0 = x_0/n_0$ . A conditional exact 100\*(1- $\alpha$ )% confidence interval

will be obtained by using the SAS procedure PROC FREQ with statement EXACT OR.

##### 4.2.4.2 Risk difference and 100\*(1- $\alpha$ )% confidence interval

For an investigational drug group with  $n_1$  subjects at risk, independent from the control group (e.g. placebo or comparator) with  $n_0$  subjects at risk, of whom  $x_1$  and  $x_0$  experience a certain event, the risk difference is estimated as  $p_1 - p_0$  with  $p_1 = x_1/n_1$  and  $p_0 = x_0/n_0$ .

Exact unconditional confidence limits for the risk difference will be obtained with SAS procedure PROC FREQ and option RISKDIFF in the TABLES statement, specifying the RISKDIFF option also in the EXACT statement.

#### 4.2.5 Exposure adjusted incidence rate and related risk estimates

##### 4.2.5.1 Exposure adjusted incidence rate and 100\*(1- $\alpha$ )% confidence interval

It will be assumed that for each of  $n$  subjects in a clinical trial the time  $t_j$  ( $j=1, \dots, n$ ) to the first occurrence of a certain event is observed, or if the event was not experienced, the (censored) time to the end of the observation period. The sequence of first occurrences of an event will be modeled to follow approximately a Poisson process with constant intensity  $\theta$ . The rate

parameter  $\theta$  will be estimated as  $\lambda = D/T$ , where  $T = \sum_{j=1}^n t_j$  and  $D$  is the number of subjects with

at least one event. Conditionally on  $T$ , an exact 100\*(1- $\alpha$ )% confidence interval for a Poisson variable with parameter  $\theta T$  and observed value  $D$  can be obtained based on ([Garwood, 1936](#)),

from which an exact  $100*(1-\alpha)\%$  confidence interval for D/T will be derived as follows ([Sahai, 1993](#); [Ulm, 1990](#)):

$$\text{Lower confidence limit } L = \frac{0.5c_{\alpha/2,2D}}{T} \text{ for } D>0, 0 \text{ otherwise,}$$

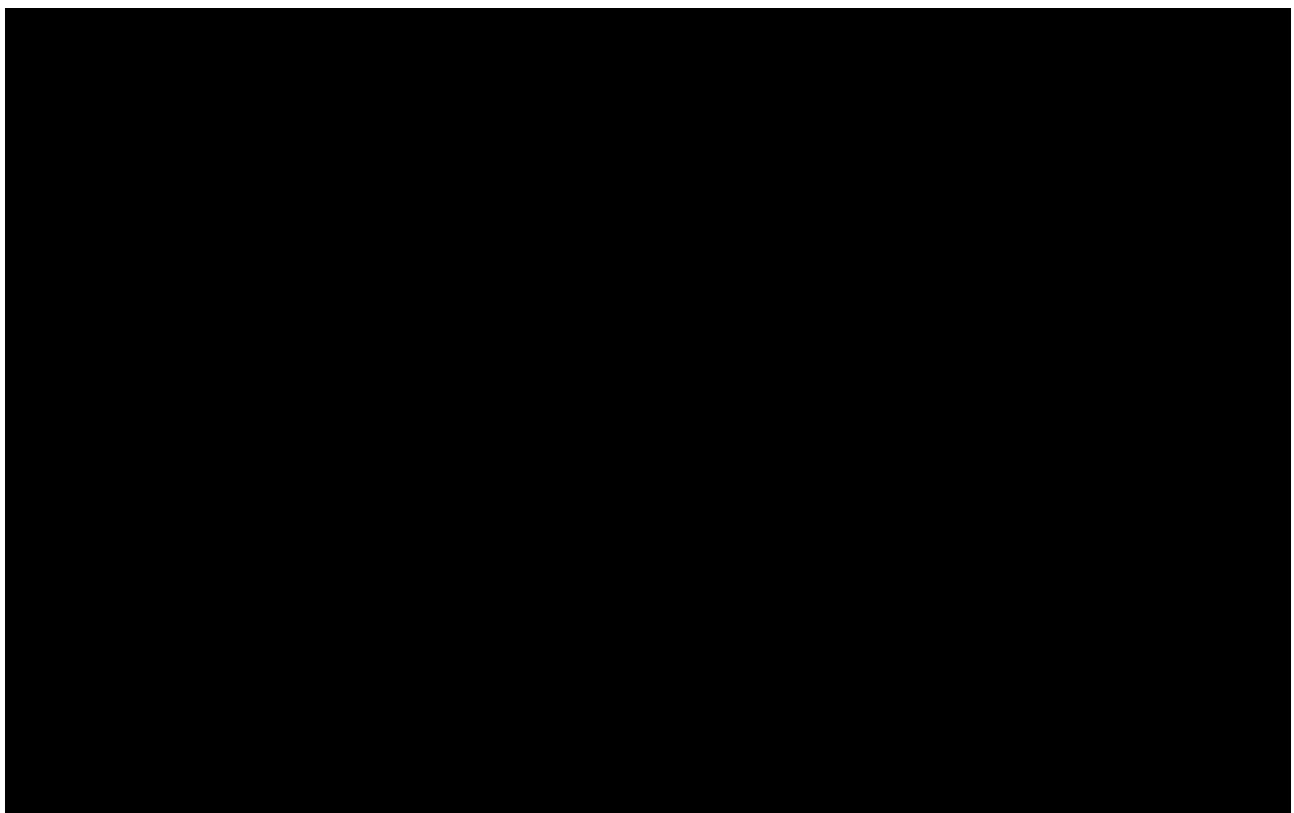
$$\text{Upper confidence limit } U = \frac{0.5c_{1-\alpha/2,2D+2}}{T}$$

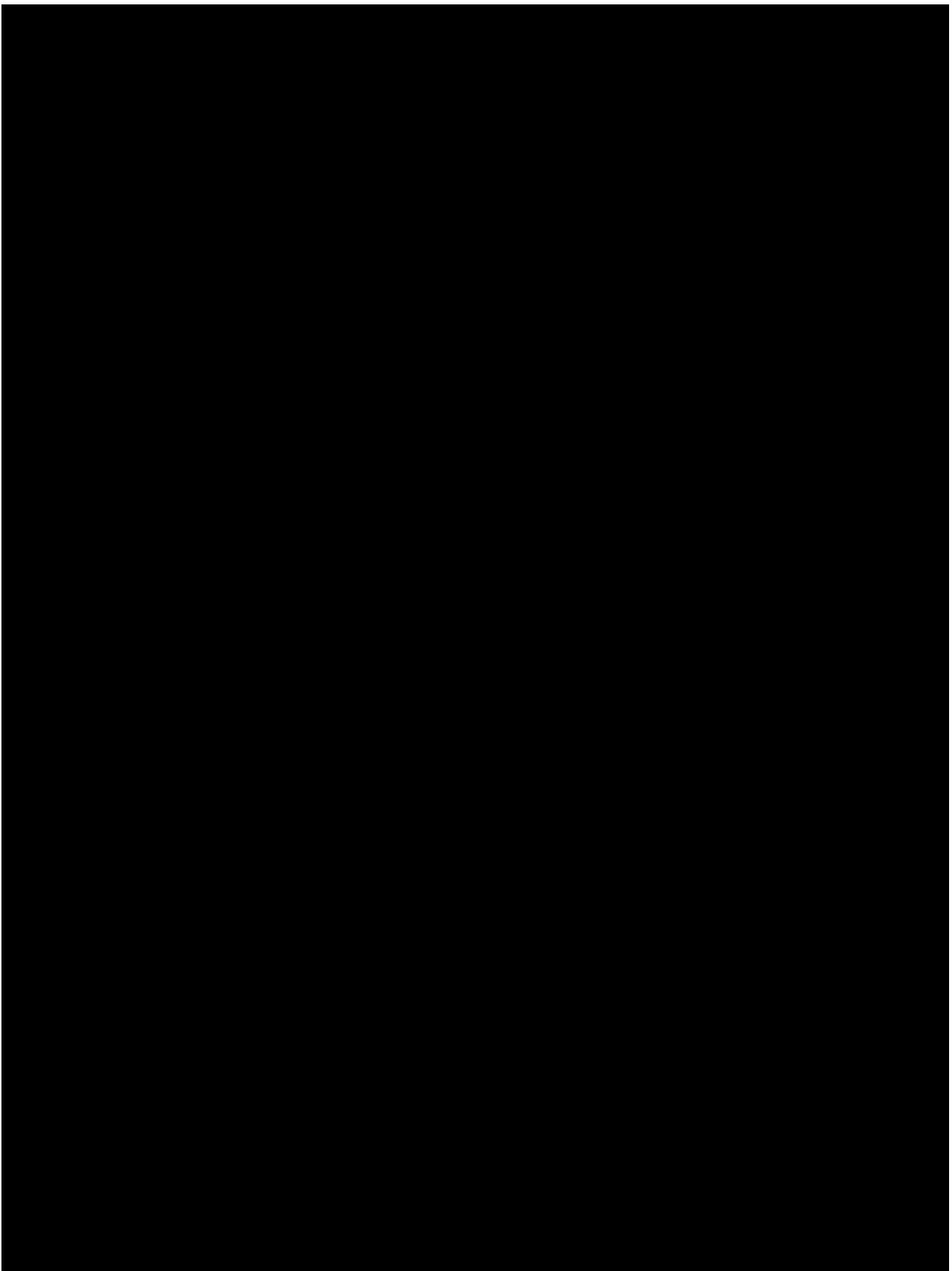
where  $c_{\alpha,k}$  is the  $\alpha$ th quantile of the Chi-square distribution with  $k$  degrees of freedom.

The example below shows how this should be handled for cases where subjects switch treatment. In particular for summarizing ‘Any AIN’ as a group, one should take into consideration the sequence of treatments while calculating exposure time for subjects.

**Table 4-3 Examples for calculating exposure time for incidence rates (IR)**

1st treatment	1st exposure	2nd treatment	2nd exposure	Event days (in terms of study day)	Exposure for IR
Placebo	100 days	150 mg	200 days	50 (1st trt) 110 (10 days into 2nd trt)	Placebo: 50 days (event) 150 mg: 10 days (event) Any AIN: 10 days (event)





#### 4.2.7 Details for Total vdH-mTSS analysis

##### 4.2.7.1 Linear Extrapolation (LE)

For the linear extrapolation consider the following variables:

- BLDay : day BL visit occurred
- W16Day/W24Day: day week 16/24 visit occurred
- W16Day/W24Day - BLDay: number of days between week 16/24 visit and baseline visit
- BL16/BL24: baseline score based on imputation algorithm using week 16/24 data.
- Wxx change from BLxx\_imp: week xx change from baseline following application of imputation algorithm using week xx data

Linear Extrapolation from W16 to W24 formula is as follows:

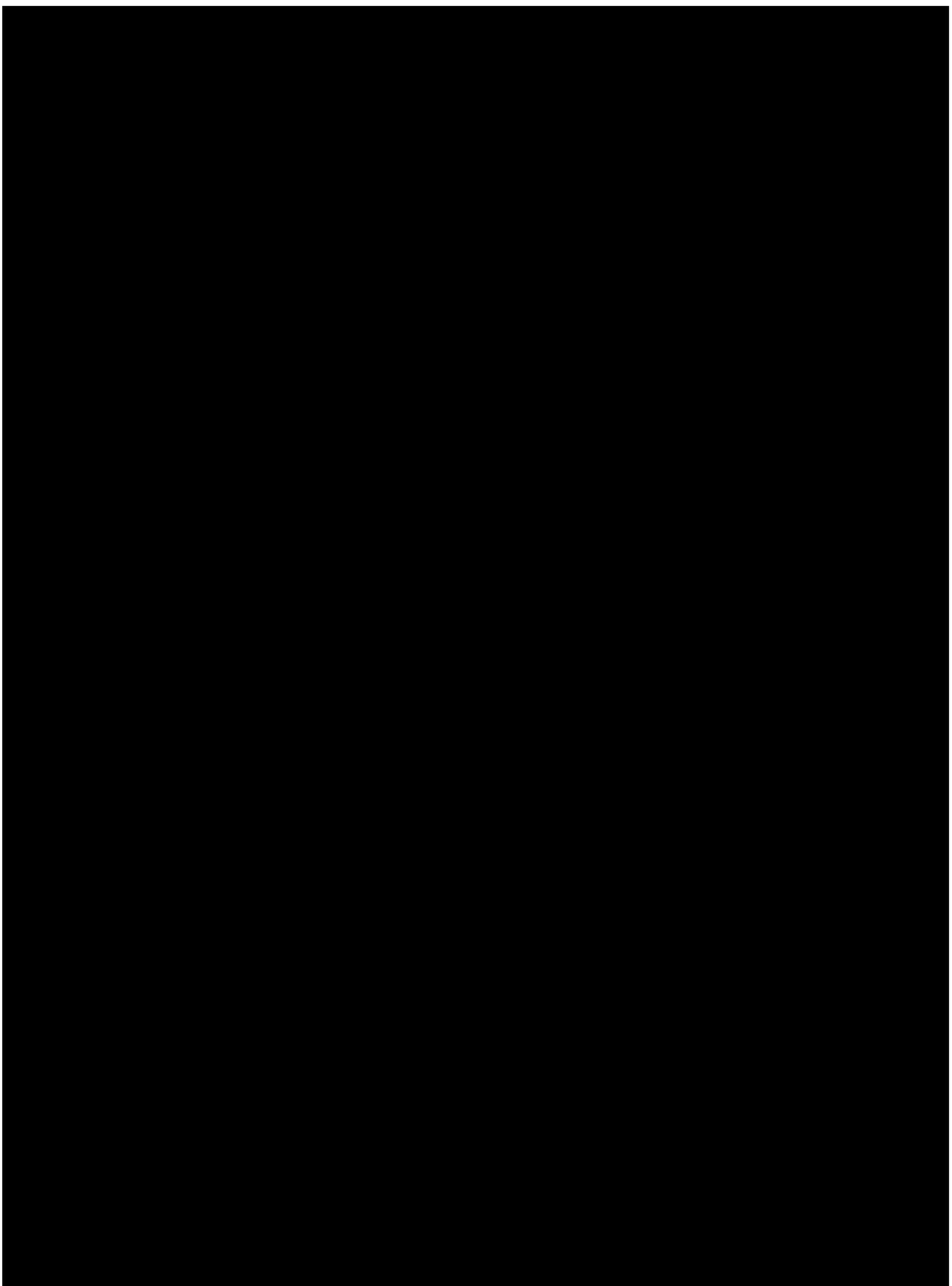
$$1. \quad W24 \text{ change from } BL\_imp = \frac{W16 \text{ change from } BL16}{W16 \text{ Day} - BL \text{ Day}} * (W24 \text{ Day} - BL \text{ Day});$$

Note: W24 Day is the assessment day of week 24 visit; if missing, use week 24 target day 169.

2.  $BL24\_imp = BL24 \text{ if available; otherwise } = BL16$ ;
3.  $W24\_imp = BL24\_imp + W24 \text{ change from } BL\_imp$ .
4. If  $W24\_imp < 0$  then set to 0 or if  $W24\_imp > \text{max attainable score}$  replace by maximum score.

For primary analysis (LE Method 1), placebo non-responders (rescued at W16) will be considered as missing at week 24 and LE will be used to extrapolate week 24 data from week 16 as for any other missing week 24 data.

For sensitivity analysis (LE Method 2), all observed week 24 data will be used. Only missing week 24 data will be imputed from LE.





## 5 Reference

AIN457A efficacy MAP M3, available in Cabinets//CREDI Projects/A/AIN457A /Administrative files/CIS (Clinical Information Sciences)/Biostatistics

Alberti KGMM, Eckel RH, Grundy SM, et al (2009) Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; american heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. *Circulation*; 120: 1640-1645.

Bishop YMM, Fienberg SE, Holland PW (1991) Discrete multivariate analysis. Cambridge MA, 11<sup>th</sup> ed. The MIT Press

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

Clopper CJ, Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26; 404–413.

Coates LC, Fransen J, Helliwell PS (2010) Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. *Ann Rheum Dis*.69(1):48-53.

Committee for medicinal products for human use (CHMP), European Medicines Agency for the Evaluation of Medicines for Human Use. (2004) Guideline on clinical investigation of medicinal products indicated for the treatment of psoriasis. CHMP/EWP/2454/02 corr document. London, UK.

CSR template available in the CREDI template area: Cabinets/CREDI Templates /CTD

Fredriksson T, Pettersson U (1978) Severe psoriasis—oral therapy with a new retinoid. *Dermatologica*; 157:238–44.

Garwood, F (1936). Fiducial limits for the Poisson distribution. *Biometrika*, 46; 441–453.

Gottlieb A, Griffiths CEM, Ho VC, et al (2005) Efficacy and tolerability of oral pimecrolimus in the treatment of moderate to severe chronic plaque type psoriasis: a double-blind, multicentre, randomized dose-finding trial. *Br J Dermatol*; 152:1219-27.

Helliwell PS, Fitzgerald O, Fransen J et al (2012) The development of candidate composite disease activity and responder indices for psoriatic arthritis (GRACE project). *Ann Rheum Dis* 2012 ;0:1–6. doi:10.1136/annrheumdis-2012-201341

Kenward, M. and Roger, J. (1997). Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. *Biometrics*; 53: 983-997.

Koch GG, Tangen CM, Jung JW, et al. (1998) Issues for covariance analysis of dichotomous and ordered categorical data from randomized clinical trials and non-parametric strategies for addressing them. *Statistics in Medicine*; 17:1863-92.

Lachin JM (2000) The assessment of relative risks . New York: Wiley

Leonardi CL, Powers JL, Matheson RT, et al (2003) Etanercept as monotherapy in patients with psoriasis. *N Engl J Med* 349:2014-22.

Leonardi CL, Kimball AB, Papp KA, et al. (2008) Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 76-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 1). *Lancet*; 371:1665-74.

Little, R.J.A and Rubin, D.B. (2002). Statistical Analysis with Missing Data. Wiley Series in Probability and Statistics, Chapter 10.

Lubeck, DP. (2004) Patient-reported outcomes and their role in the assessment of rheumatoid arthritis. *Pharmacoconomics* 22(1): 27-38.

McInnes, I.B. et al. (2013) Efficacy and safety of ustekinumab in patients with active psoriatic arthritis: 1 year results of the phase 3, multicentre, double-blind, placebo-controlled PSUMMIT 1 trial. *Lancet* . ISSN 1474-547X (doi:10.1016/S0140-6736(13)60594-2)

Menter A, Gottlieb A, Feldman S, et al (2008) Guidelines of care for the management of psoriasis and psoriatic arthritis – Section 1. Overview of psoriasis and guidelines of care for the treatment of psoriasis with biologics. *J Am Acad Dermatol*; 58:826-850.

Newcombe RG (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine*; 17: 857-872.

Ostergaard, Mikkel et al (2009). The OMERACT Psoriatic Arthritis Magnetic Resonance Imaging Scoring System (PsAMRIS): Definitions of Key Pathologies, Suggested MRI Sequences, and Preliminary Scoring System for PsA Hands. *The Journal of Rheumatology* 36:8 1816-1824.

Papp KA, Tyring S, Lahfa M et al (2005) A global phase II randomized controlled trial of etanercept in psoriasis: safety, efficacy, and effect of dose reduction. *Br J Dermatol*; 152:1304-12.

Papp KA, Langley RG, Lebwohl M, et al. (2008) Efficacy and safety of ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with psoriasis: 52-week results from a randomised, double-blind, placebo-controlled trial (PHOENIX 2). *Lancet*; 371:1675-84

Compound Case Retrieval Strategy available in Cabinets/CREDI Projects/A/AIN457A/Integrated Medical Safety

Ritchlin C, Rahman P, Kavanaugh A, et al. (2013). Efficacy and safety of the anti-IL-12/23 p40 monoclonal antibody, ustekinumab, in patients with active psoriatic arthritis despite conventional non-biological and biological anti-tumour necrosis factor therapy: 6-month and 1-year results of the phase 3, multicentre, double-blind, placebo controlled, randomised PSUMMIT 2 trial. *Ann Rheum Dis*. 2014 Jan 30. doi: 10.1136/annrheumdis-2013-204655. [Epub ahead of print]

Rubin, D.B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Safety Profiling Plan stored in CREDI (Cabinets/CREDI Projects/A/AIN457A/Integrated Medical Safety)

Sahai H, Khurshid Anwer (1993). Confidence intervals for the mean of a poisson distribution: a review. *Biom J*, 35 (7); 857-867

Sampson HA, Muñoz-Furlong A, Bock SA, et al (2005) Symposium on the Definition and Management of Anaphylaxis: Summary report. *J Allergy Clin Immunol* 2005;115:584-91.

Sampson HA, Muñoz-Furlong A, Campbell RL, et al (2006) Second Symposium on the Definition and Management of Anaphylaxis: Summary Report - Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network Symposium. *Ann Emerg Med*;47(4):373-80.

Schaefer, J.L. (1997). *Analysis of Incomplete Multivariate Data*, Chapman&Hall.

Tyring S, Gottlieb A, Papp KA, et al (2006) Etanercept and clinical outcomes, fatigue, and depression in psoriasis: double-blind placebo-controlled randomised phase III trial. *Lancet* 367:29-35.

Ulm K (1990). A simple method to calculate the confidence interval of a standard mortality ratio. *American Journal of Epidemiology*, 131(2); 373-375

Van der Heijde DM (1999). How to Read Radiographs According to Sharp/van der Heijde Method. *J Rheumatol*. 26(3):743-745.

Van der Heijde D, Sharp J, Wassenberg S, Gladman DD (2005) Psoriatic arthritis imaging: a review of scoring methods. *Ann Rheum Dis*; 64: Suppl 2:ii 61-64.

Weisman S, Pollack CR, Gottschalk RW (2003) Psoriasis disease severity measures: comparing efficacy of treatments for severe psoriasis. *J Dermatolog Treat*; 14: 158-165.

Zink RC and Koch GG (2012). NParCov3: A SAS/IML Macro for Nonparametric Randomization-Based Analysis of Covariance. *Journal of Statistical Software*; 50(3): 1-17.