NOVARTIS

Clinical Development

CAIN457/Secukinumab/Cosentyx®

Clinical Trial Protocol CAIN457F2366 / NCT02745080

A randomized, double-blind, active control, multicenter study to evaluate the efficacy at week 52 of subcutaneously administered secukinumab monotherapy compared with subcutaneously administered adalimumab monotherapy in patients with active psoriatic arthritis

Statistical Analysis Plan (SAP) Amendment 2

Author: Trial Statistician, [redacted]

Document type: SAP Documentation

Document status: Final

Release date: 15-Oct-2019

Number of pages: 56

Document History – Changes compared to previous final version of SAP

Date	Reason for update	Outcome for update	Section and title impacted (Current)
10-April-2017	First release	Final version	
[redacted]			
11-Jul-2019	Early treatment discontinuation visit	Change premature study treatment discontinuation visit from before week 52 to before week 50 (last dosing visit)	Section 2.5.1
11-Jul-2019	cDMARDs list	Add Apremilast to cDMARDs list	Section 2.4.2
[redacted]			
11-Jul-2019	MI process detail	Provide more details on the process of implementing Multiple Imputation (MI)	Section 2.5.3
11-Jul-2019	MI code for resolution of enthesitis	Add SAS code for implementing Multiple Imputation for resolution of enthesitis, which is not derived from components.	Section 6.2.3
[redacted]			

Date	Reason for update	Outcome for update	Section and title impacted (Current)
15-Oct-2019	Clarified code of proc mi	Added more details of SAS code	Appendix 6.2.2.4
[redacted]			
15-Oct-2019	Clarified the text about study treatment discontinuation	Clarified that study treatment discontinuation is before or at Week 50	Sections 2.6.1 and 2.6.5

Table of contents

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	ClASsification criteria for Psoriatic ARthritis
cDMARD	conventional Disease modifying anti-rheumatic drugs (also known as nonbiologic DMARDs)
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
[redacted]	[redacted]
DMARD	Disease Modifying Antirheumatic Drug
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	Electronic Case Report/Record Form
[redacted]	[redacted]
ESR	Erythrocyte Sedimentation Rate
[redacted]	[redacted]
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
[redacted]	[redacted]
hsCRP	High sensitivity C-Reaction Protein
[redacted]	[redacted]
IL	Interleukin

i.v.	Intravenous(ly)
IVR	Interactive Voice Response
IWR	Interactive Web Response
[redacted]	[redacted]
LDL	Low Density Lipoprotein
[redacted]	[redacted]
LLN	Lower Limit Normal
LLOQ	Lower Limit of Quantification
MAR	Missing at Random
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
MedDRA	Medical Dictionary for Drug Regulatory Affairs
[redacted]	[redacted]
mmHg	Millimeter mercury
MMRM	Mixed-effects Repeated Measures Model
[redacted]	[redacted]
MTX	Methotrexate
NSAID	Non-steroidal andti-inflammatory drug
[redacted]	[redacted]
PASI	Psoriatic Area and Severity Index
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
PoC	Proof of Concept
PRO	Patient-reported Outcomes
PsA	Psoriatic Arthritis
PsAQoL	Psoriatic Arthritis Quality of Life questionnaire
[redacted]	[redacted]
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
[redacted]	[redacted]
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	TNFα Inhibitor Inadequate Responders
ULN	Upper Limit of Normal
VAS	Visual Analog Scale
WBC	White Blood Cells
WHO	World Health Organization
[redacted]	[redacted]

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 is a randomized, double-blind, active control, multicenter, parallel-group trial evaluating secukinumab monotherapy and adalimumab monotherapy in approximately 850 subjects with active PsA. A screening period up to 8 weeks before randomization will assess subject eligibility.Efficacy assessments will occur through Study Week 52. Two follow-up visits at Week 60 and 68 will occur thereafter. Total maximum study duration, including screening period, is up to 76 weeks.

At Baseline, subjects whose eligibility is confirmed will be randomized to either Group 1 (secukinumab 300 mg as 2 x 150 mg s.c. injections), or Group 2 (adalimumab 40 mg as 1 s.c. injection) with approximately 425 subjects/group. In order to maintain the blind, both groups will receive 1 or 2 placebo s.c. injections to keep consistency in the number of injections at each dosing visit. Secukinumab is available in 150 mg/1.0 mL pre-filled syringes (PFS) and adalimumab is available in 0.4 mL PFS. Placebo (1.0 and 0.5 mL, PFS) is also available.

Group 1 - Secukinumab 300 mg s.c.

Secukinumab 300 mg (2 x 1 mL PFS) will be administered at Baseline, Weeks 1, 2, 3 and 4, followed by dosing every 4 weeks until Week 48. In addition, all Group 1 subjects will receive placebo (1 x 1mL PFS) at given visits in order to maintain the blind.

Group 2 - Adalimumab 40 mg s.c.

Adalimumab 40 mg (1 x 0.4 mL PFS) will be administered at Baseline followed by dosing every 2 weeks until Week 50. In addition, all Group 2 subjects will receive placebo (1 x 0.5mL or 2 x 0.5mL PFS) at given visits in order to maintain the blind.

1.2 Study objectives and endpoints

1.2.1 Primary objective

To demonstrate that the efficacy of secukinumab monotherapy 300 mg s.c. at Week 52 is superior to adalimumab monotherapy (40 mg s.c.) based on the proportion of subjects achieving an American College of Rheumatology 20 (ACR20) response.

1.2.2 Secondary objectives

To demonstrate that:

1. Secukinumab monotherapy (300 mg s.c.) is superior to adalimumab monotherapy (40 mg s.c.) at Week 52, based on the proportion of subjects achieving PASI90 response.
2. Secukinumab monotherapy (300 mg s.c.) is superior to adalimumab monotherapy (40 mg s.c.) at Week 52, based on the proportion of subjects achieving an ACR50 response.

3. The improvement (change) from baseline on secukinumab monotherapy (300 mg s.c.) is superior to adalimumab monotherapy (40 mg s.c.) at Week 52, for the Health Assessment Questionnaire – Disability Index (HAQ-DI©) score.
4. Secukinumab monotherapy (300 mg s.c.) is superior to adalimumab monotherapy (40 mg s.c.) at Week 52, based on the proportion of subjects achieving the resolution of enthesitis based on LEI.
5. An additional secondary objective is to evaluate the safety and tolerability of secukinumab monotherapy (300 mg s.c.) compared with adalimumab monotherapy (40 mg s.c.) as assessed by vital signs, clinical laboratory values, and adverse events monitoring.

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 comparison of secukinumab 300mg with adalimumab will be performed at Week 52 unless otherwise specified.

Efficacy and safety data for the entire treatment period will be presented by treatment groups.

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. Two subsets of FAS are defined as follows:

- Enthesitis subset (LEI): The enthesitis subset will include all FAS subjects who have enthesitis based on LEI at baseline.
- [redacted]
- Psoriasis subset: The psoriasis subset will include all FAS subject who have >= 3% of the body surface area (BSA) affected by psoriatic skin involvement at baseline.

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 Treatment groups

The summaries by treatment will be performed by the randomized treatment.

Randomized treatment:

- AIN457 300 mg
- Adalimumab 40 mg

[redacted]

[redacted]

2.4 Patient disposition, demographics and other baseline characteristics

2.4.1 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 the treatment period for each treatment group and all subjects.

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

2.4.2 Background and demographic characteristics

The following background and demographic variables 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)^2

Categorical variables:

- Age categories (<65 years, 65 years and older, 75 years and older)
- Gender
- Race
- Ethnicity

The following disease specific baseline characteristics will be summarized by treatment groups:

- ACR components and other disease-related measures (e.g. [redacted], presence of enthesitis (based on LEI [redacted]), [redacted], time since first diagnosis of psoriatic arthritis, [redacted], subjects with psoriasis ≥ 3%)

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.

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

2.5.1 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, ≥ 1 week, ≥ 2 weeks, ≥ 3 weeks, ≥ 4 weeks, ≥ 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 minimum of (last dose of the treatment + 84 days) and (last visit date).

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.

2.5.2 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 therapies 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). The number and percentage of subjects receiving the following three categories of prior or concomitant medications will be summarized separately in one table: cDMARDs, MTX and non-MTX cDMARDs. This is the list of cDMARDs: Sulfasalazine, Leflunomide, Cyclosporine, Methotrexate, azathioprine, minocycline and apremilast.

Prohibited medication is defined in section 5.5.8 of the protocol.

The number and rates of subjects taking Prohibited medication will be presented.

2.6 Analysis of the primary objective

2.6.1 Primary endpoint

As the Primary objective of the study is to compare the efficacy of secukinumab **monotherapy** 300 mg s.c. with adalimumab **monotherapy** (40 mg s.c.) with respect to ACR20 at Week 52, the estimand of interest is the ACR20 response at week 52 in monotherapy patients who do not permanently discontinue study medication prematurely nor require rescue cDMARDS. The estimand will be evaluated in the FAS population by comparing the odds of response to each treatment (odds ratio). The primary efficacy endpoint is a composite endpoint. Clinical response, failure to permanently discontinue study medication prematurely, and lack of need for rescue medication will be required for a patient to achieve treatment success. The primary endpoint is the proportion of patients with **monotherapy** ACR20 response at Week 52 where **monotherapy** ACR20 response is defined as meeting the following three conditions:

- achieving American College of Rheumatology 20 (ACR20) response
- no permanent study treatment (Secukinumab or adalimumab) discontinuation before or at Week 50 (the last dosing visit)
- no use of cDMARDs (including MTX) after week 36 (regardless of the starting time of taking cDMARDs)

A patient who does not meet any of the three conditions above is regarded as a monotherapy non-responder.

The list of possible cDMARDs for criterion 3 are the following: sulfasalazine, leflunomide, cyclosporine, methotrexate, azathioprine, minocycline and apremilast.

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

2.6.2 Statistical hypothesis, model, and method of analysis

The statistical hypothesis for monotherapy ACR20 is that there is no difference in the proportion of subjects fulfilling the monotherapy ACR20 criteria at Week 52 in the secukinumab 300 mg group vs. adalimumab.

Let p_j denote the proportion of monotherapy ACR20 responders at Week 52 for treatment regimens j, j=0, 1 where

0 corresponds to adalimumab,

1 corresponds to secukinumab 300 mg

In statistical terms, H_1: p1 = p0, H_{1A}: p1 $\neq$ p0, i.e.

H_1: secukinumab is not different to adalimumab regimen for signs and symptoms (monotherapy ACR20 response) at Week 52.

The primary endpoint of monotherapy ACR20 at Week 52 in the FAS will be evaluated using a logistic regression model with treatment as a factor and weight as a covariate. Odds ratios will

be computed for comparisons of secukinumab 300 mg vs. adalimumab regimen utilizing the logistic regression model fitted.

2.6.3 Handling of missing values/censoring/discontinuations

Missing data

Patients who were unblinded prior to the scheduled timepoint will be considered non-responders from the time of unblinding up to Week 52.

Missing data for ACR20 response and other binary efficacy variables (e.g. monotherapy response of ACR50, [redacted], PASI90, etc.) for data up to Week 52 will be handled by Multiple Imputation (MI) if a subject neither discontinues the study treatment prematurely nor takes cDMARDs after week 36 but still has missing value for ACR20 (e.g., misses some visits). Details of the implementation is in Section 6.2.2.4.

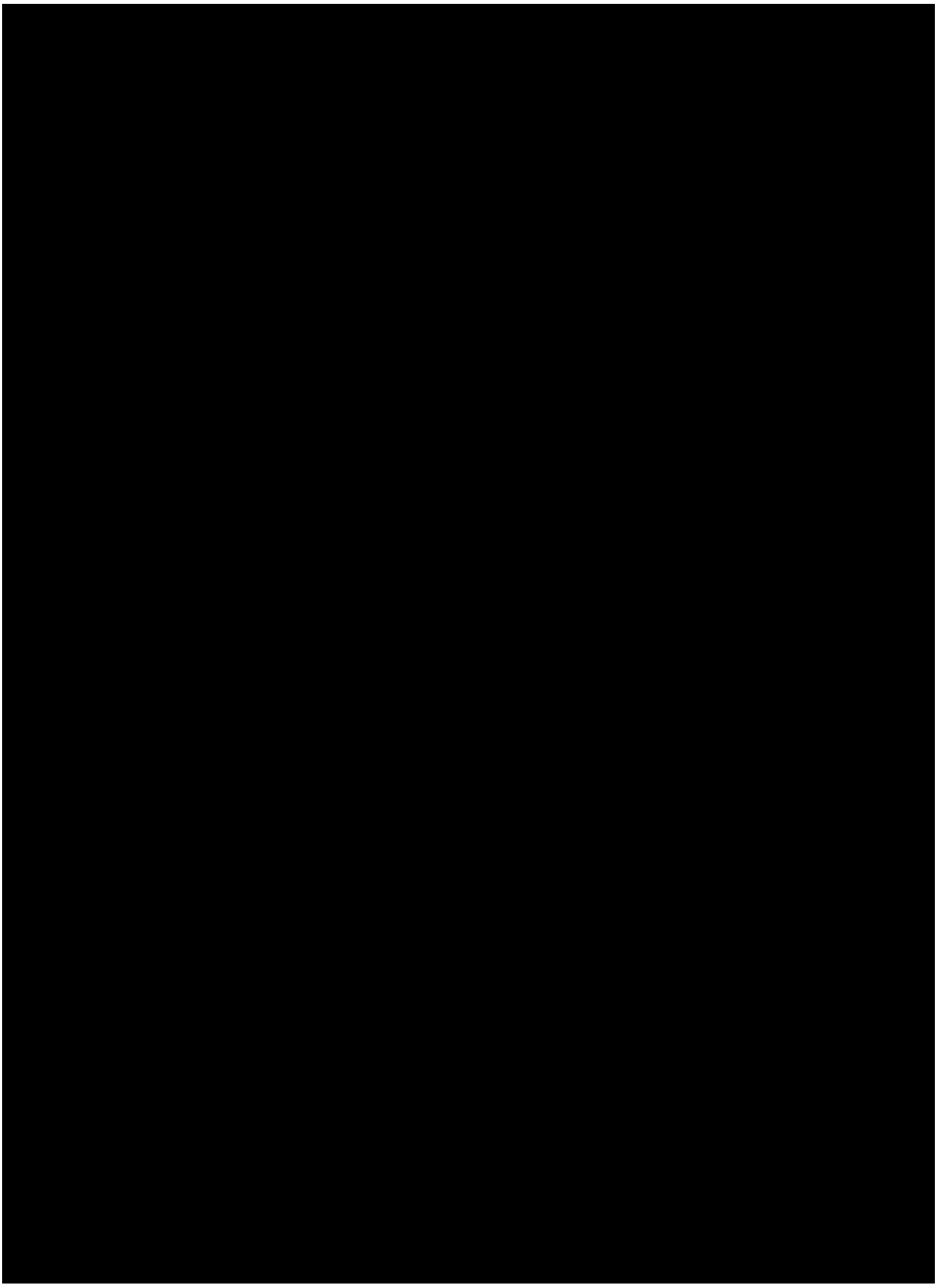
Multiple imputation (MI) is a simulation based approach where missing values are replaced by multiple Bayesian draws from the conditional distribution of missing data given the observed data and covariates, creating multiple completed data sets. These completed data sets can then be analyzed using standard methods. Within this analysis the ACR composites or PASI score will be imputed and response variables will be derived based on the imputed scores. In the multiple imputation analysis the response status will be imputed based on the individual treatment arm information.

Missing data for continuous efficacy endpoints due to missed or skipped visits will not be applied MI. Continuous variables (e.g. ACR20 components) will be analyzed using a mixed-effects repeated measures model (MMRM) which is valid for addressing missing value under the missing at random (MAR) assumption. As such, single-point imputation of missing data will not be performed (e.g. LOCF) nor employing MI. Details of the implementation is in Section 6.2.1.2.

For analysis based on subset of subjects (i.e. Psoriasis, Enthesitis [redacted]), a subject with missing baseline values will be excluded in the analysis.

For use of cDMARDs (including MTX) after Week 36, the start date and end date of cDMARDS use will be compared to the upper limit of the Week 36 analysis visit window (day 274). For partial or missing dates the following imputation rule will be used:

- If only date is missing (month and year are present), then for Concomitant Medication start date, use the first day of the month, for end date, use the last day of the month.
- If both month and date are missing, then for Concomitant Medication start date, use January 1st of the year, for end date, use December 31st of the year.
- If year, month and date are all missing, then for Concomitant Medication start date, no imputation will be applied but we assume the start date is before day 274. For Concomitant Medication end date, no imputation will be applied but we assume the end date is ongoing.



2.7 Analysis of secondary objectives

Except the HAQ-DI© score, the secondary efficacy variables are composite endpoints defined in a similar fashion as the primary endpoint and are listed below. Secondary efficacy variables will be analyzed using the FAS population unless otherwise specified.

- **Monotherapy** PASI90 response at Week 52 in Superiority test for secukinumab 300 mg based on Psoriasis subset
- **Monotherapy** ACR50 response at Week 52 in Superiority test for secukinumab 300 mg
- Change from baseline in HAQ-DI© score at Week 52 in Superiority test for secukinumab 300 mg based on a MMRM model.
- **Monotherapy** resolution of enthesitis based on LEI at Week 52 in Superiority test for secukinumab 300 mg based on Enthesitis subset (LEI).

All the monotherapy responses (binary endpoints) in the list of the secondary endpoints (PASI90, ACR50, resolution of enthesitis) are defined as meeting the following three conditions:

1. achieving the corresponding clinical response as defined after Figure 2-1
2. no permanent study treatment (Secukinumab or adalimumab) discontinuation before Week 52
3. no use of cDMARDs (including MTX) after week 36 (regardless of taking time of taking cDMARDs).

For HAQ-DI© score, if a subject does not meet any of the conditions 2 and 3 above, then the data of this score after the time of taking cDMARDs (including MTX) or premature discontinuation of study treatment will be set to missing (censored) and analyzed using MMRM. For analyses of HAQ-DI© score using MMRM, if baseline value is missing then this subject will be removed from the, 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.

Testing strategy to control type I error

The following hypotheses will be included in the hierarchical testing strategy, and type-I-errors will be set such that family-wise two-sided type-I-error of 5% is kept. The inferential testing procedure will only continue if the previous test was rejected at two-sided 5% level.

Primary objective

- H_1: secukinumab 300 mg is not different to adalimumab regimen with respect to monotherapy ACR20 response at Week 52.

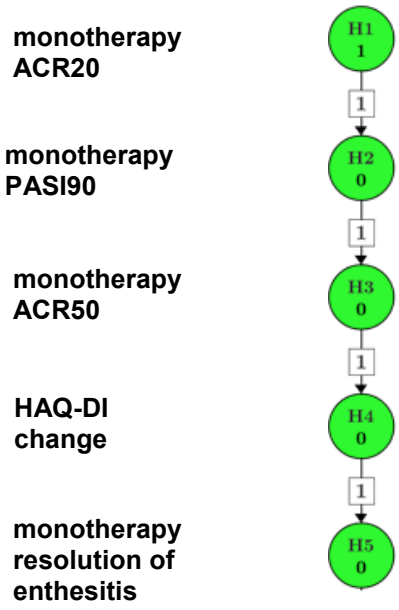
Secondary objectives

The same method of imputation for binary endpoints that is being used for the primary is also being used for secondary endpoints

- H_2: secukinumab 300 mg is not different to adalimumab regimen with respect to monotherapy PASI90 response at Week 52
- H_3: secukinumab 300 mg is not different to adalimumab regimen with respect to monotherapy ACR50 response at Week 52
- H_4: secukinumab 300 mg is not different to adalimumab regimen with respect to the change from baseline in HAQ-DI© at Week 52
- H_5: secukinumab 300 mg is not different to adalimumab regimen with respect to monotherapy resolution of enthesitis based on LEI at Week 52

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 rate will be set to two-sided α=5% and it will be controlled with the proposed hierarchical testing strategy. If H1 is rejected at α, each of the following hypothesies will be tested at α. If H2 is rejected at α, H3 will be tested at α and so on.

Monotherapy ACR50 at Week 52

Monotherapy response at Week 52 to ACR50 in the FAS will be evaluated using a logistic regression model with treatment as a factor and weight as a covariate.

Monotherapy PASI90 response

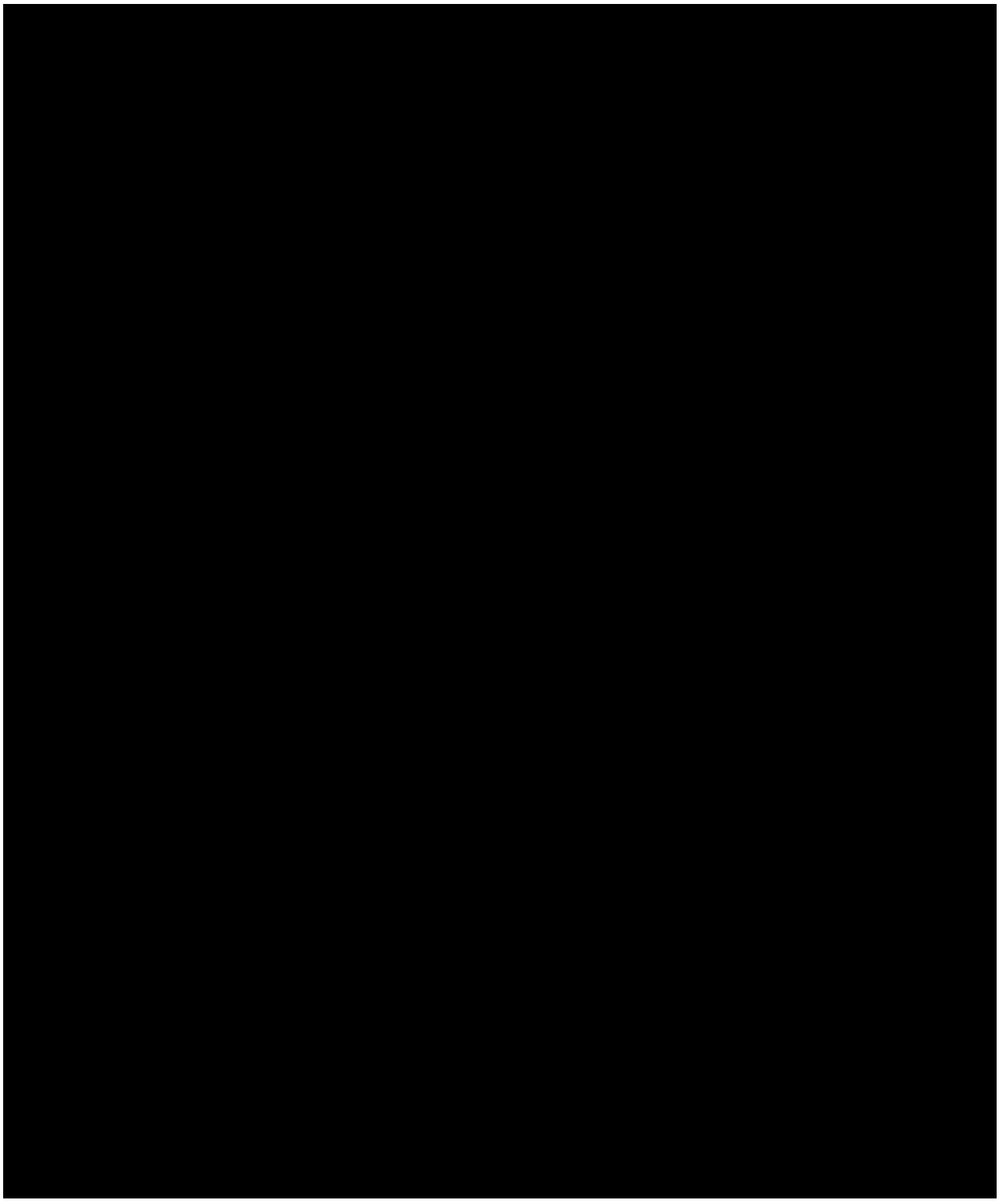
Monotherapy PASI90 response at Week 52 will be evaluated using a logistic regression model with treatment as a factor and weight as a covariate based on Psoriasis subset.

Physical function (HAQ-DI©)

Between-treatment differences in the change in HAQ-DI© will be evaluated using a MMRM with treatment regimen, analysis visit as factors, and weight and baseline HAQ-DI© score as continuous covariates. Treatment by analysis visit and baseline 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 regimen at different analysis visits will be determined from the pairwise comparisons performed between secukinumab regimen and adalimumab at the appropriate analysis visits.

Monotherapy Resolution of enthesitis

Monotherapy resolution of enthesitis based on LEI at Week 52 will be evaluated using a logistic regression model with treatment as a factor and weight as a covariate based on Enthesitis subset (LEI).



2.9 Safety analyses

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.

2.9.1 Adverse events (AEs)

The crude incidence of treatment emergent adverse events (TEAEs) (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 5.2.4.1. Graphical displays of the crude incidence rates will be presented for all AEs and serious AEs by system organ class.

Adverse events reported will be presented first by alphabetical order of system organ class and then in descending frequency according to its incidence in secukinumab group starting from the most common event. For ties in incidence rate of secukinumab group, sort adverse events by incidence rate of adalimumab group. 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 Programming Dataset Specification (PDS).

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 the 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	Risk	Notables for (vitals/ECG), lab criteria
Entire treatment	Cruide incidence	Cruide incidence	Cruide incidence	Cruide incidence	Cruide incidence

2.9.2 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” 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:

change from baseline = post baseline value – baseline value

For each parameter, the maximum change (maximum decrease and maximum increase) from baseline, if appropriate for each study phase, 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 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 for study level reports and the pooled summary of clinical safety (SCS).

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

CTCAE v4.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
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; >1.5 - 3.0 x ULN	>3.0 baseline; >3.0 - 6.0 x ULN	>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
Glucose decreased (Hypoglycemia)	<LLN - 3.0 mmol/L	<3.0 - 2.2 mmol/L	<2.2 - 1.7 mmol/L	<1.7 mmol/L
Cholesterol high	>ULN - 7.75 mmol/L	>7.75 - 10.34 mmol/L	>10.34 - 12.92 mmol/L	>12.92 mmol/L
Hypertriglyceridemia	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 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 **(Hy's Law**) 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 *pure hepatocellular injury*. This does not mean that cases of ALT or AST >3xULN & TBL >2xULN & ALP ≥2xULN may not result in severe DILI.

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, three individual subject data listings will be provided for liver events, laboratory values and viral serology, autoimmunity, imaging, pathology, drug abuse for patients with liver events.

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 ">")."

2.9.3 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.

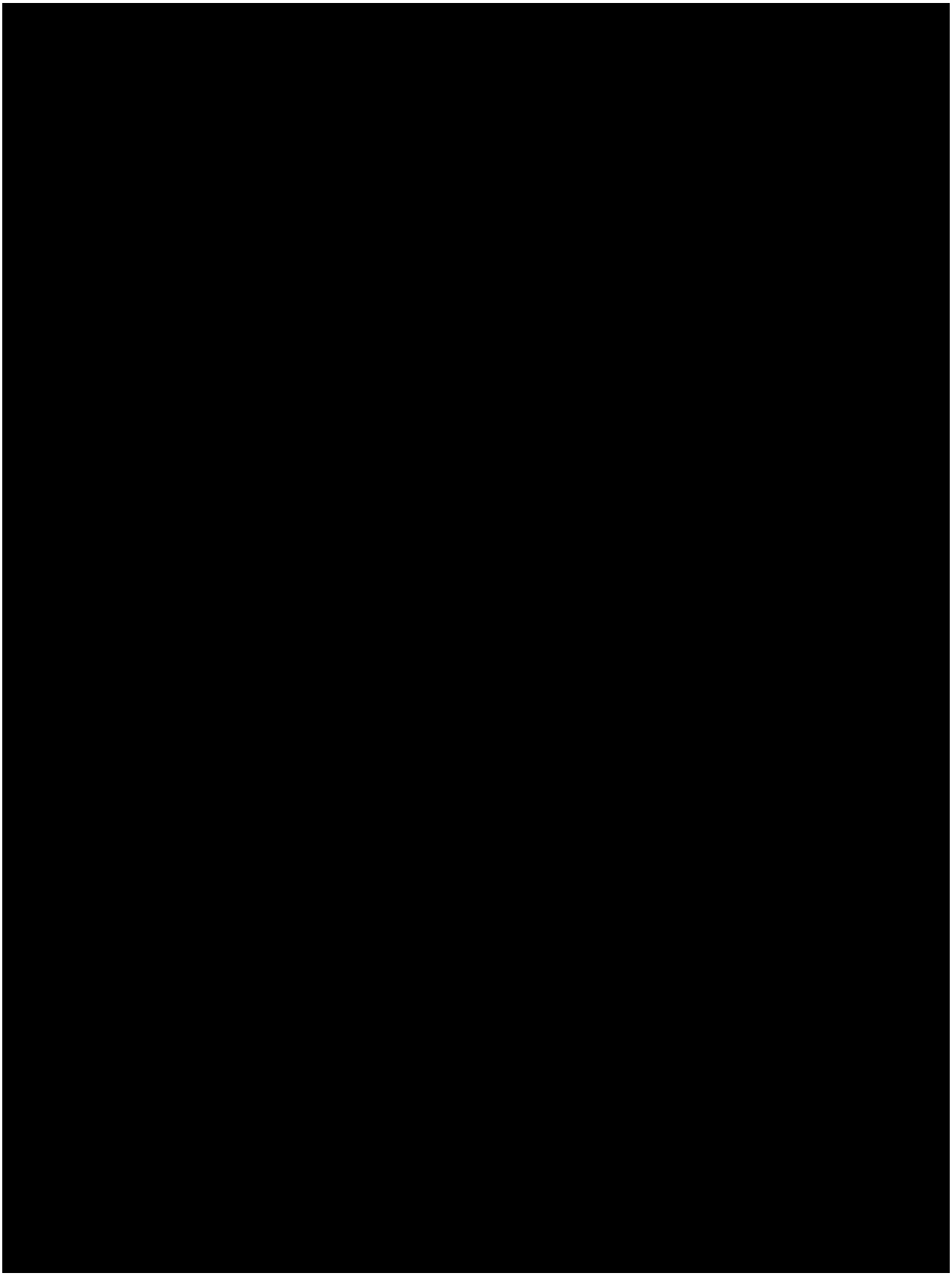
Analysis in vital sign measurement using descriptive summary statistics for the change from baseline for each post-baseline visit will be performed. These descriptive summaries will be presented by vital sign and treatment group. Change from baseline will only be summarized for subjects with both baseline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

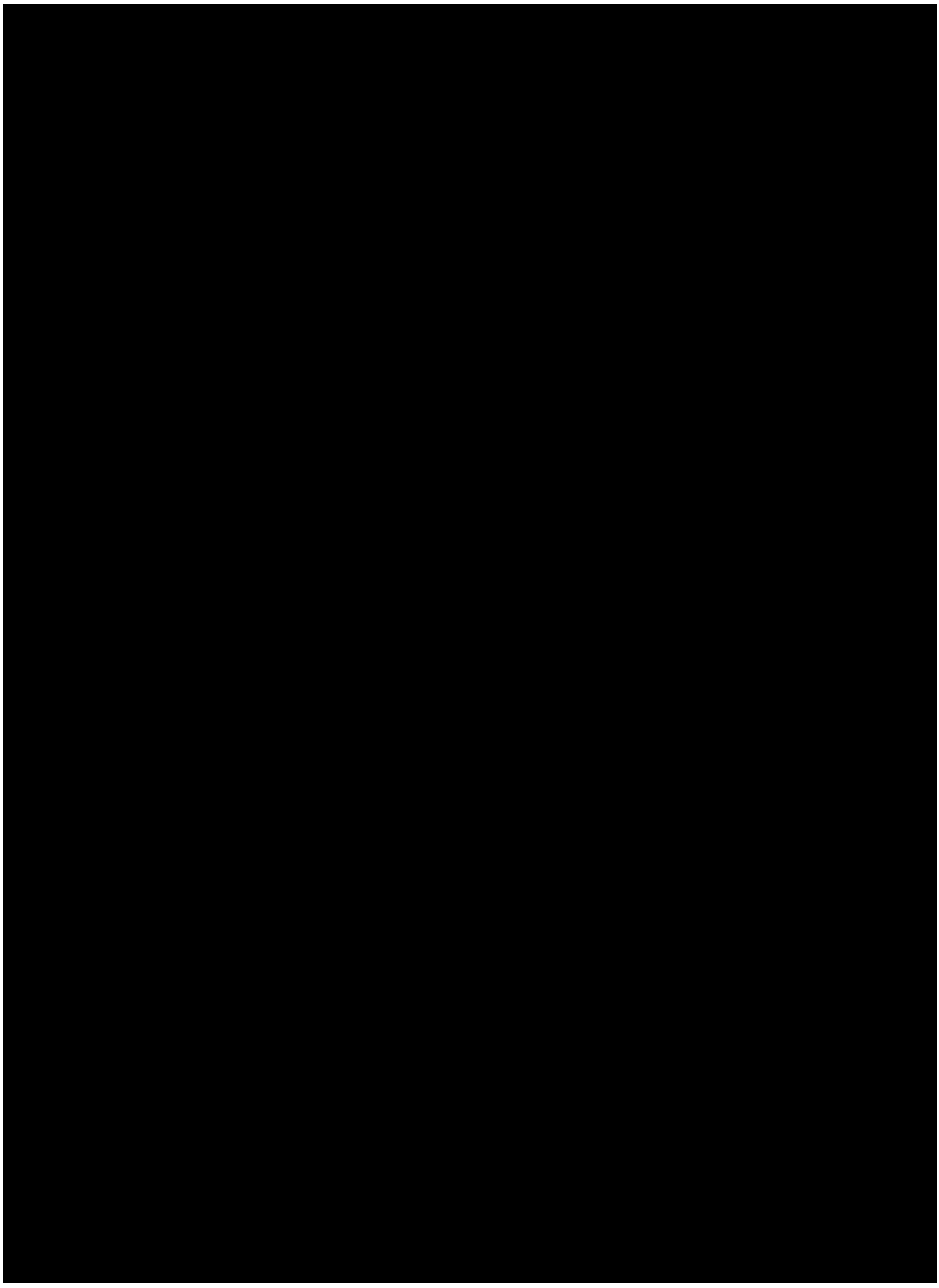
change from baseline = post-baseline value – baseline value

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)	>= 140 mmHg or < 90 mmHg
Diastolic blood pressure (mmHg)	>=90 mmHg or <60 mmHg
Pulse (bpm)	> 100 bpm or <60 bpm





[redacted]

2.12.3 Efficacy Evaluation

2.12.3.1 Description of efficacy variables

ACR 20/50/[redacted]

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., ≥):

- 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 PsA disease activity (VAS 100 mm)
  - Physician's global assessment of PsA disease activity (VAS 100 mm)
  - Patient self-assessed disability (Health Assessment Questionnaire [HAQ©] 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 52. Primarily, CRP will be used to calculate ACR response: ESR will only be used in the event CRP is missing.

ACR50 [redacted] 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, [redacted] 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 = 3rd 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 PsA 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©] 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.

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 that 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. "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.

There are two definitions for **HAQ-DI response** that will be used for analysis:

Definition 1: an improvement of at least 0.3 score points compared to baseline.

Definition 2: an improvement of at least 0.35 score points compared to baseline.

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.

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 PsA 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.

Erythrocyte sedimentation rate (ESR)

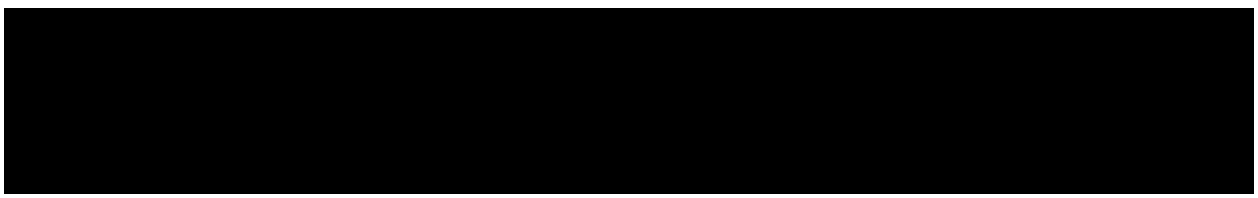
Blood for ESR, which is helpful in diagnosing inflammatory diseases and is used to monitor disease activity and response to therapy, will be obtained at scheduled visits.

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.

Presence of enthesitis

If enthesitis based on LEI 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.



[redacted]

Resolution of enthesitis

Resolution of enthesitis (based on LEI [redacted]) is defined as absence of enthesitis at any post-baseline visit in subjects included in the Enthesitis subset (LEI [redacted]).

[redacted]

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. 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.

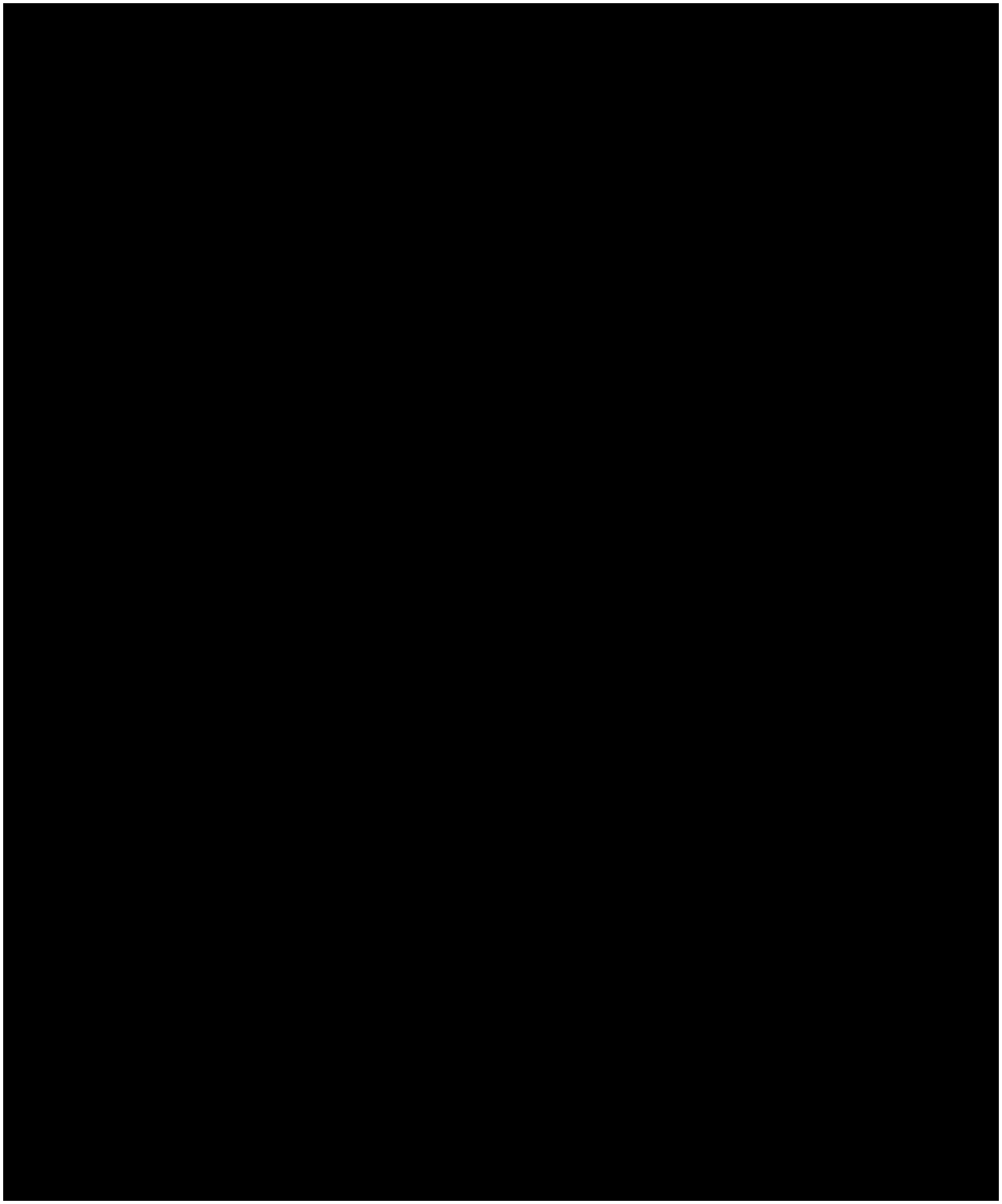
****Buttocks are assessed as part of the Lower limbs (L) body region.

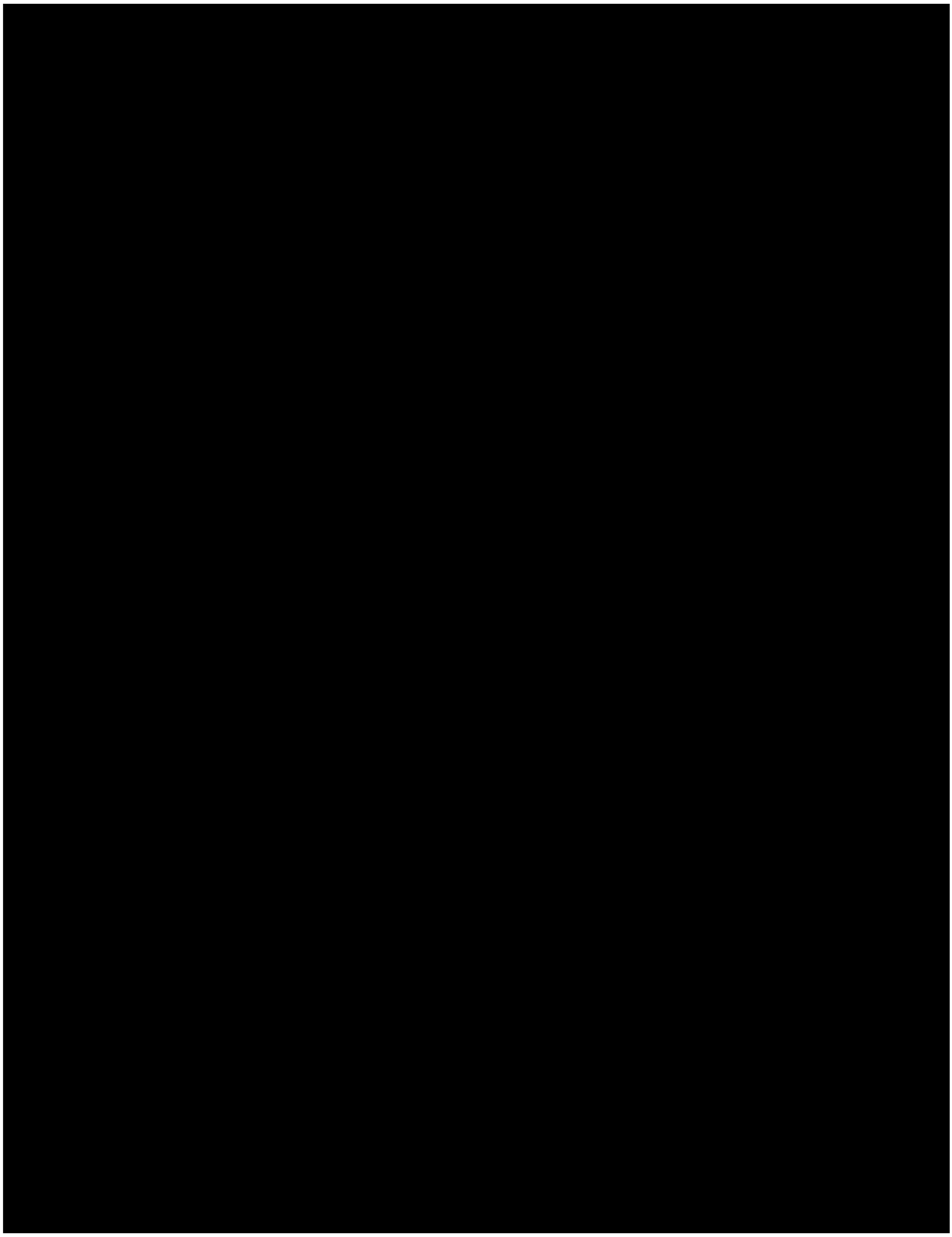
Because the head and neck, upper limbs, trunk and lower limbs correspond to approximately 10%, 20%, 30% and 40% of the body surface area, respectively, the PASI score is calculated using the formula:

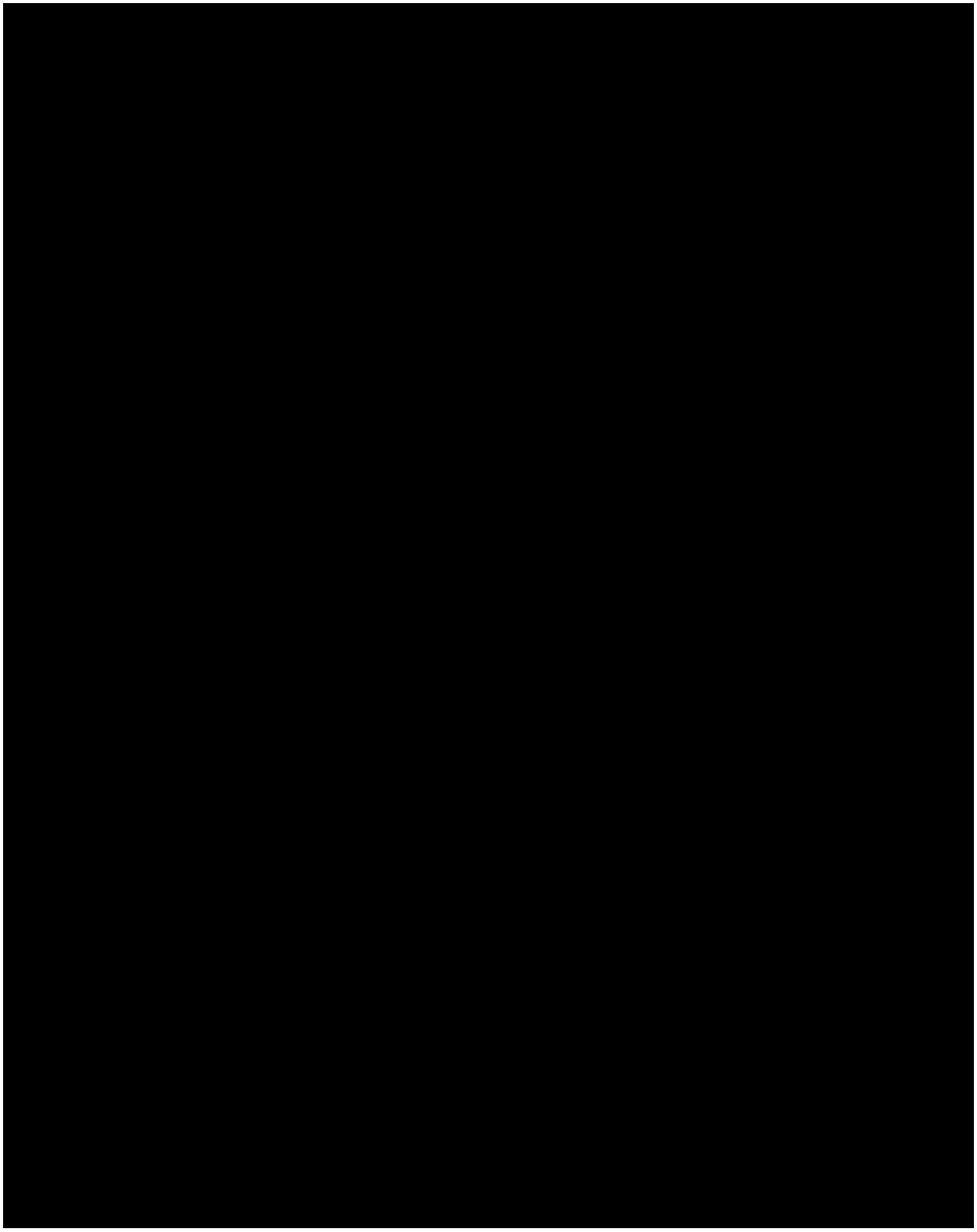
$PASI = 0.1(E_H+I_H+D_H)A_H + 0.2(E_U+I_U+D_U)A_U + 0.3(E_T+I_T+D_T)A_T + 0.4(E_L+I_L+D_L)A_L$

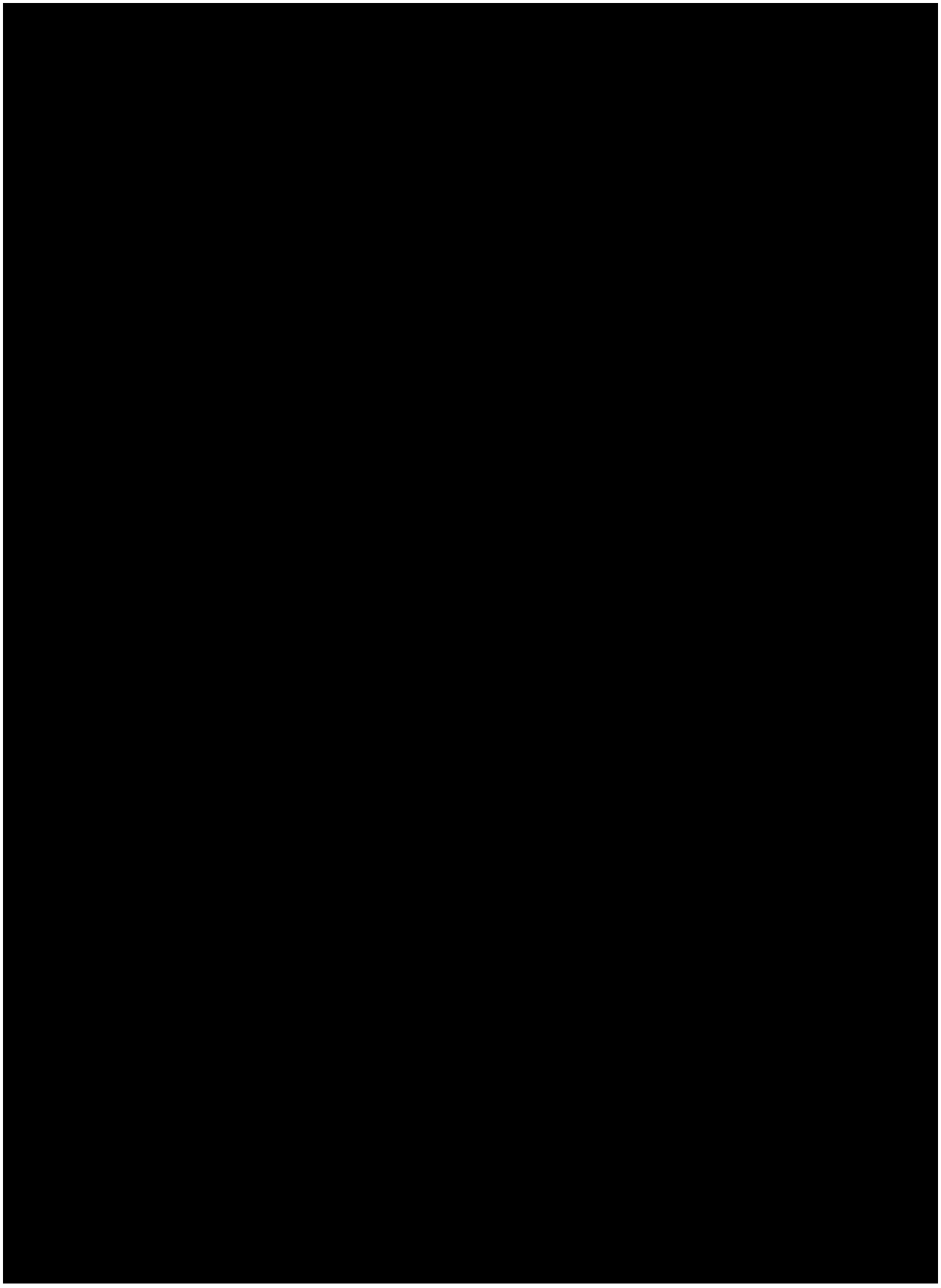
The keys for the letters are provided in Table 2-7.

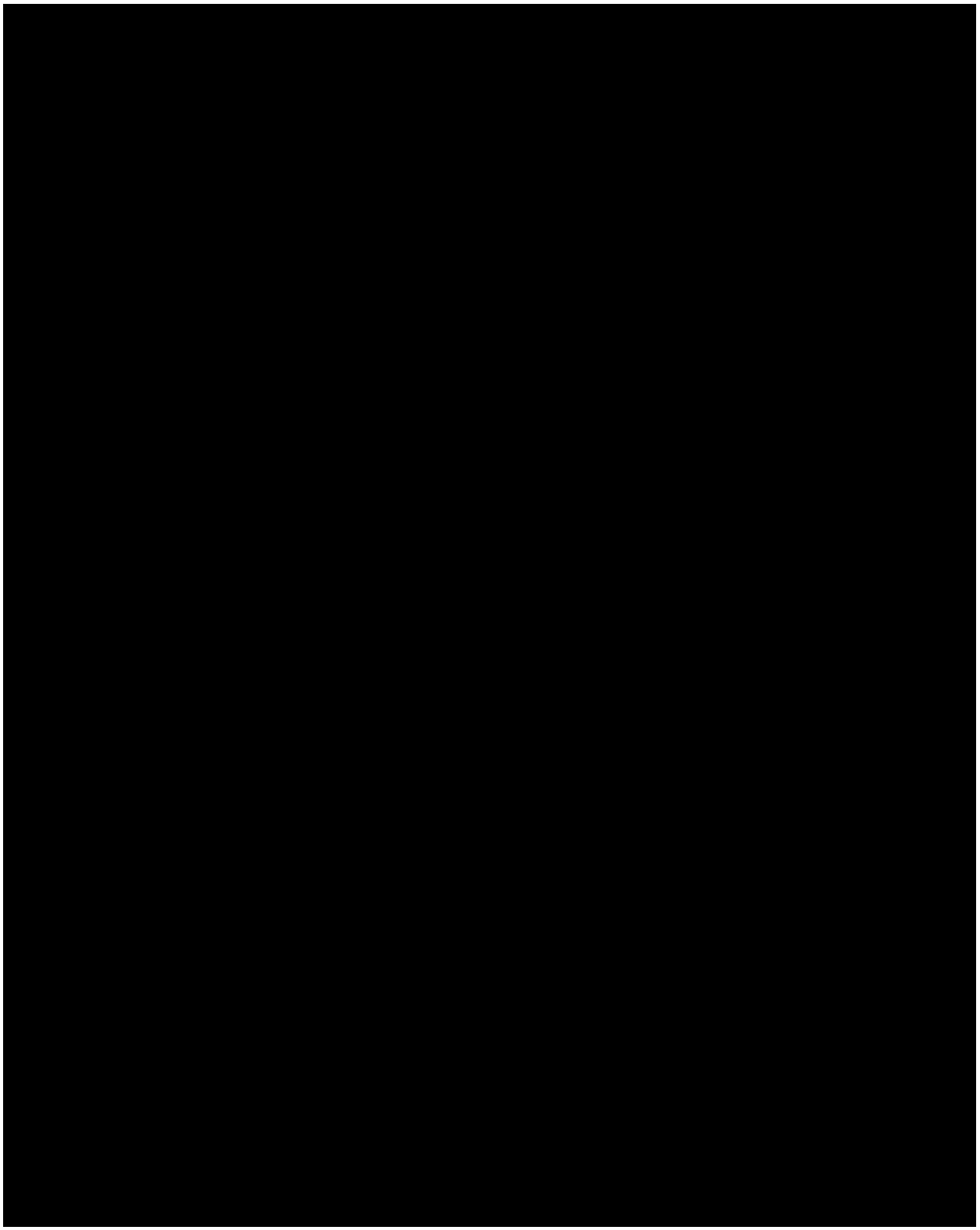
PASI scores can range from a lower value of 0, corresponding to no signs of psoriasis, up to a theoretic maximum of 72.0. The total score comes from eCRF.











3 Sample size calculation

A family-wise error will be set to two-sided α=0.05 to control for type I error. The primary objective of secukinumab 300 mg versus adalimumab with respect to ACR50ACR20 response at Week 52 will be tested at a two sided 0.05 alpha level.

An ACR20 response rate of about 50% for the TNFα inhibitor naïve without MTX use (monotherapy) population at week 48 was reported in Humira study (Gladman 2007). The response on secukinumab 300 mg is estimated to be 62% in the TNFα inhibitor naïve monotherapy population based on the lower bound of 80% confidence interval of Meta-analysis results of the study FUTURE2 (McInnes 2015) and unpublished study FUTURE3. With 425 subjects per treatment group there would be approximately 94% power to detect a treatment difference of 12% at a two-sided 0.05 alpha level in ACR20 response rates (Chi Square test, NQuery 7.0) between secukinumab 300 mg and adalimumab in the evaluation of the primary efficacy hypothesis at Week 52. The overall sample size will be 850 patient for a randomization ratio of 1:1 (secukinumab 300mg 425 patients, and adalimumab 425 patient). The estimated power with the chosen sample size for other efficacy endpoints based on the data available for adalimumab and meta-analysis results of the study FUTURE2 and study FUTURE3 are summarized in Table 3-2..

Table 3-1 Assumption for sample size and power calculation

Endpoint	Secukinumab 300 mg expected values	Adalimumab (N=76) observed values
ACR20	62%	50%
PASI90	63%	38% (N=40)
ACR50	49%	38%
HAQ-DI change*	-0.59 (0.5)	-0.4 (0.5)(N=61)
Resolution of enthesitis	61%	33%

*mean (standard deviation)

Sources: The value for Secukinumab is from the lower bound of 80% confidence interval of a meta-analysis based on FUTURE2 and unpublished results from FUTURE3 data, the values for other endpoints for Secukinumab are from the means of the same meta-analysis, the values for Adalimumab are from Humira® study (Gladman 2007).

Based on Table 3-1, power calculation for superiority tests of different endpoints are presented below in Table 3-2.

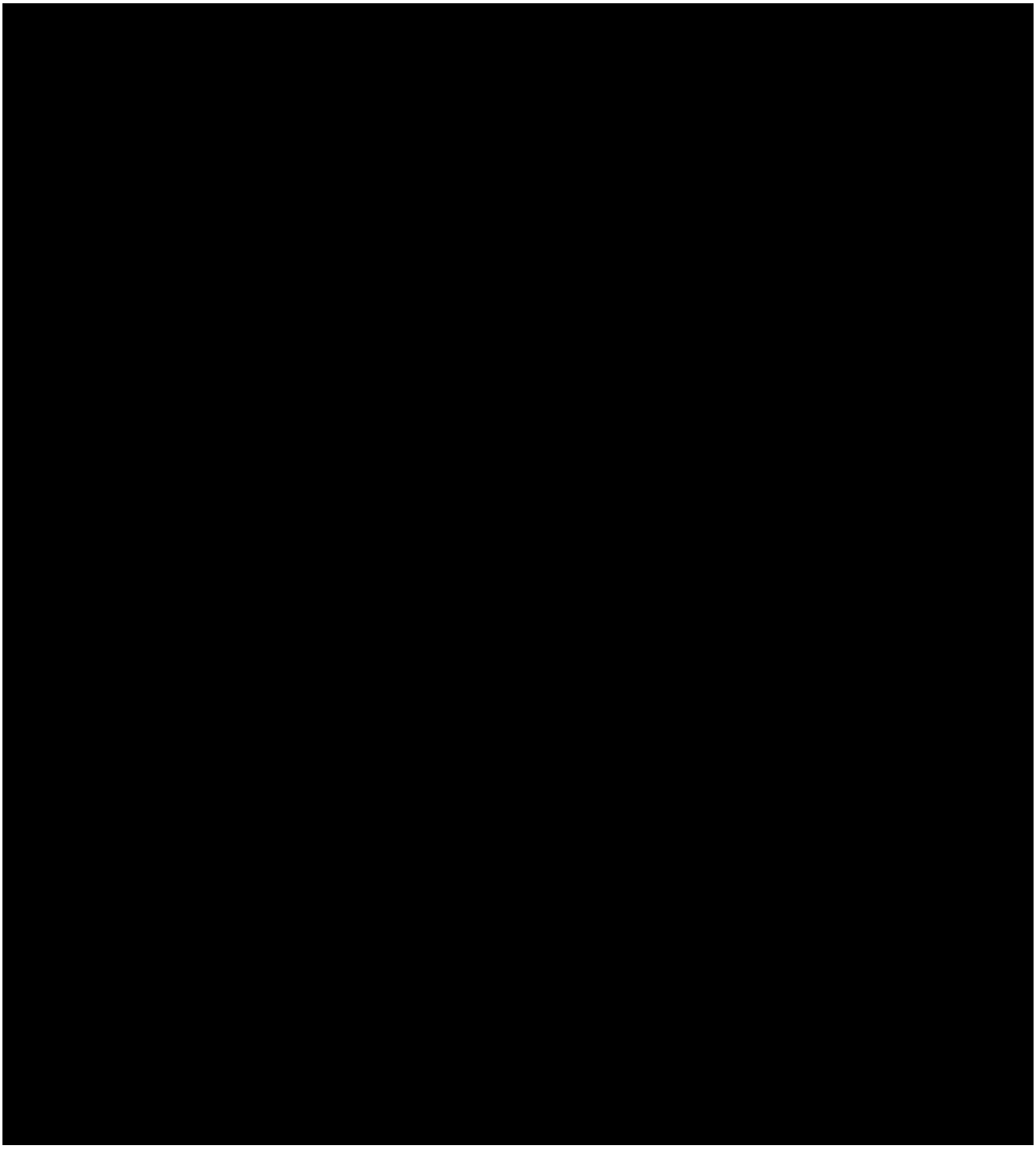
Table 3-2 Power for superiority tests for primary and secondary endpoints (two-sided, α=0.05)

Endpoint	Power*
ACR20 (H1)	94%
PASI90 (H2)	99%
ACR50 (H3)	91%
HAQ-DI (H4)	99%
Resolution of enthesitis (H5)	99%

* power does not consider dependence of the endpoints in the Testing Strategy, power for PASI90 and enthesitis is based on the assumption that a subset of half of the population (213 subjects per group) has data of these two endpoints.

4 Change to protocol specified analyses

All the analysis specified in the protocol for CSR is detailed in this SAP.



6 Appendix

6.1 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 scheduled 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).

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 (W50) 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 6-1 Analysis Visit Windows

Analysis Visit	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Baseline	≤1*	≤1*	≤1*	≤1*	≤1*	≤1*	≤1*	≤1*
Week 2	2 – 26				2 - 26			
Week 4	27 – 50	2 - 134		2 - 71	27 - 50			2-50
Week 8	51 – 78				51 - 78			51 – 78
Week 12	79 – 106		2 - 148	72 - 106	79 - 106	2 - 148		79 – 106
Week 16	107 – 134			107 - 155	107 - 155			107 – 134
Week 20	135 – 162							135 – 162
Week 24	163 – 190	135 - 316	149 - 316	156 - 211	156 - 211	149 - 211	2 - 316	163 – 190
Week 28	191 – 218							191 – 218
Week 32	219 – 246			212 - 267	212 - 267	212 - 267		219 – 246
Week 36	247 – 274							247 – 274
Week 40	275 – 302			268 - 344	268 - 344	268 - 344		275 – 302
Week 44	303 – 330							303 – 330
Week 48	331 – 358							331 – 358
Week 52	359 – 498	317 - 449	317 - 498	345 - 498	345 - 498	345 - 498	317 - 449	359 – 498
Week 68		450 - 498					450 - 498	

* The first administration of randomized study treatment (first dose) is defined as Day 1.

Group1: ACR components, ESR, High sensitivity C-Reactive protein

Group2: [redacted]

Group3: Hematology, blood chemistry, urinalysis, Lipids, [redacted]

Group4: [redacted]

Group5: LEI, [redacted], PASI, [redacted]

Group6: Urine pregnancy test

Group7: [redacted]

Group8: Vital signs

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 = 0.25*(Week number of current visit)* 7 + 0.75 * (Week number of next applicable visit)*7 + 1. Lower limit of the first applicable visit is always Day 2.

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. Taking into account the fact that the protocol stipulation is "the study treatment should not be administered less than 14 days from the previous administration" and bi-weekly dosing schedule for Adalimumab, which leads to all visit delays accumulative, the following rule will be used to determine the appropriate measurement to be summarized in the event of multiple measurements within a visit window.

Table 6-2 Rules for flagging variables

Timing of measurement	Type of data	Rule
Baseline	All data	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. If there are multiple assessments on Day 1, following rules will apply: (a) If assessment time exists, - select the last available measurement prior to reference start date/time considering time; - if no measurement prior to reference start date/time considering time, select the earliest measurement post reference start date/time considering time . (b) If assessment time does not exist, select the available measurement from the lowest CRF visit number.
Post-baseline efficacy	All data	• **The last assessment within the window will be used.** • Cases where the same parameter is recorded more than once on the same date will be handled as follows: o If time of completion exists the earliest measurement will be used; o If time does not exist the measurement from the lowest CRF visit number will be used.
Post-baseline safety	Summary visit information (e.g. lab)	• **The last assessment within the window will be used.** • • Cases where the same parameter is recorded more than once on the same date will be handled as follows: o If time of completion exists the earliest measurement will be used; o If time does not exist the measurement from the lowest CRF visit number will be used.
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.

6.2 Statistical methodology and assumptions

6.2.1 Analysis of continuous data

6.2.1.1 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.

6.2.1.2 Mixed-effects repeated measures model

Endpoints with continuous data type expected to be normally distributed (e.g, [redacted]) will be analyzed using a mixed-effects repeated measures model (MMRM) with treatment, 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 regimen at different analysis visits will be determined from the comparisons performed between secukinumab regimen and adalimumab at the appropriate analysis visits.

SAS code for mixed model:

```
proc mixed data=aaa;

class TRT USUBJID AVISITN;

model CHG=TRT 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.

Details of Implementation

Steps in implementing MMRM for continuous variables

1. For each continuous endpoint, set the data collected after patient discontinued study treatment to missing.
2. For each continuous endpoint, set data collected after patient took cDMARDs after Week 36 to missing.
3. Proceed with MMRM

6.2.2 Analysis of binary (and categorical) data

6.2.2.1 Summary statistics for binary and categorical data

Summary statistics for discrete variables will be presented in contingency tables and will include count and frequency in each category. If applicable, confidence intervals 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=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

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

and the upper limit is

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

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

For binary response variables (e.g. for ACR20/50/[redacted], HAQ-DI responder, PASI 75, [redacted]) 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=formmatted;

Tables response*trt/Fisher;

Run;

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

6.2.2.2 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 regimen versus control(s) utilizing the logistic regression model fitted.

SAS code for logistic regression:

```
proc genmod data = acr20 descending;
class trt_ ;
model aval = trt_  weight  / link=logit dist=bin;
lsmeans trt_ / diff;
estimate 'AIN457 300 mg vs ADA 40 mg' trt_ 1 -1;
ods output Estimates=imp_est;
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 300 mg", "ADA 40 mg");
Run;
```

6.2.2.3 Cochran-Mantel-Haenszel Test

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

6.2.2.4 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 (Week 24) using Markov Chain Monte Carlo (MCMC) method with EM algorithm.

Impute the missing values 100 times (NIMPUTE) with a seed=457<studycode> for each component necessary to calculate the final score as shown below:

```
proc mi data= out=imp min= &minval max= &maxval minmaxiter=10000000 nimpute=100
seed=4572366;
by trt;
var weight_base Baseline Week_2  Week_4  Week_8  Week_12
Week_16  Week_20  Week_24  Week_28  Week_32  Week_36
Week_40  Week_44  Week_48  Week_52;
```

```
    mcmc chain=multiple initial=em;
run;
```

where &minval and &maxval refer to minimum and maximum imputed values. For tender joint, minval=0, maxval=78; For swollen joint, minval=0, maxval=76; For VAS scores, minval=0, maxval=100; For CRP and ESR, minval=0, maxval= (no maximum value); For PASI total score, minval=0, maxval= 72; For [redacted], minval=0, maxval=10; For [redacted], minval=0.96, maxval=9.4 (Hansen et al 2017); For mCPDAI, minval=0, maxval=12; For [redacted], minval=0, maxval=164 . By defining the min and max during imputation, each variable will be confirmed to be within the range of its scale. The imputation of CRP and ESR will be based on log(value+1) and will be converted back to original value scale after imputation.

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 label1;
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 miexpres;
  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_ ;
        model aval = trt_  weight  / link=logit dist=bin;
        lsmeans trt_ / diff;
        estimate 'AIN457 300 mg vs ADA 40 mg' trt_ 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;
```

The MI code for the resolution of enthesitis [redacted] is as follows:

```
proc mi data=enthr seed=4572366 nimpute=100 out=enthr_mi;
    class  Baseline Week_2   Week_4        Week_8        Week_12       Week_16
        Week_20       Week_24       Week_28       Week_32       Week_36       Week_40
        Week_44       Week_48       Week 52;
     by trt01p;
```

```
fcs logistic(  Baseline Week_2      Week_4      Week_8      Week_12     Week_16
        Week_20     Week_24     Week_28     Week_32     Week_36     Week_40
        Week_44     Week_48     Week 52 /details) reg( weight/details);
    var  weight  Baseline Week_2  Week_4      Week_8      Week_12     Week_16
        Week_20     Week_24     Week_28     Week_32     Week_36     Week_40
        Week_44     Week_48     Week 52;
  run;
```

Details of Implementation

Steps in implementing MI and monotherapy response for binary variables (except for resolution of enthesitis [redacted]).

1. For each continuous component (or total score for PASI, [redacted]) dataset, set the data collected after patient discontinued study treatment to missing.
2. For each continuous component (or total score for PASI, [redacted]) dataset, set data collected after patient took cDMARDs after Week 36 to missing.
3. Apply multiple imputation to create multiple sets of complete dataset
4. Calculate the binary endpoint for each complete dataset.
5. Set as non-response for the visits after the End of Treatment visit (i.e. if End of Treatment happened at Week 16, all visits after Week 16 will be set as non-response).
6. Set as non-response for the visits after the visit when patient took cDMARDs after Week 36 (i.e. Study day 274, the upper limit of analysis visit window of Week 36).
7. Set as non-response for the visits after the visit of unblinding before end of study date.
8. Proceed with the analysis.

Steps in implementing MI and monotherapy response for resolution of enthesitis [redacted]:

1. Set data collected after patient discontinued study treatment to missing.
2. Set data collected after patient took cDMARDs after Week 36 to missing.
3. Apply multiple imputation to create multiple sets of complete dataset.
4. Set as non-response the visits after the End of Treatment visit (i.e. if End of Treatment happened at Week 16, all visits after Week 16 will be set as non-response)
5. Set as non-response the visits after the visit when patient took cDMARDs after Week 36 (i.e. Study day 274, the upper limit of analysis visit window of Week 36)
6. Proceed with the analysis

6.2.3 Crude incidence and related risk estimates

6.2.3.1 Crude incidence and 100*(1-α)% confidence interval

For n subjects, each at risk to experience a certain event with probability π, the crude incidence is estimated as $p=x/n$, where x is the number of subjects with the event.

Absolute and relative frequencies will be displayed as well as 95% confidence interval for the relative frequency based on the score method including continuity correction (Newcombe 1998).

With z as (1-α/2)-quantile of the standard normal distribution (SAS: z=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

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

and the upper limit is

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

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

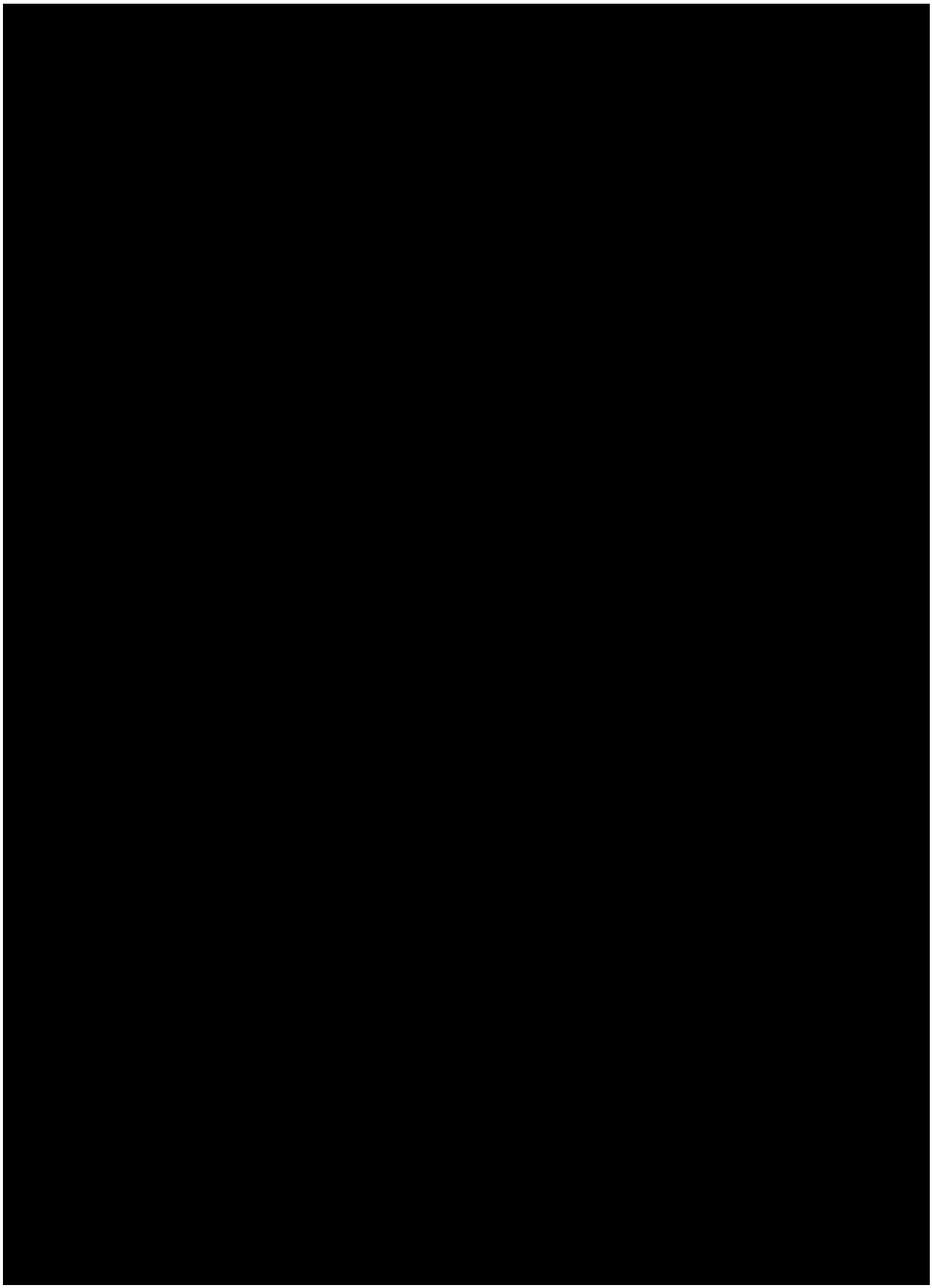
If appropriate, an exact 100*(1-α)% confidence interval (Clopper-Pearson 1934) will be obtained by using the SAS procedure PROC FREQ with the EXACT BINOMIAL statement. However, the confidence interval derived via the score method including continuity correction will be the default in safety analyses.

6.2.3.2 Odds ratio and 100*(1-α)% 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 π_1 and π_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-α)% confidence interval will be obtained by using the SAS procedure PROC FREQ with statement EXACT OR.

[redacted]



[redacted]

7 Reference

[redacted]

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

AIN457 safety MAP M3, available in Cabinets//CREDI Projects/A/AIN457A/Administrative files/CIS (Clinical Information Sciences)/Biostatistics

Biostatistical Guidance on Analysis Sets in Clinical Trials, available in Cabinets/CREDI TABULU/B&SR/CIS Process Documentation/Guidances (outside of ESOPS)/Others

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

'CIS liver safety' guidance, available in Cabinets/CREDI TABULU/B&SR/CIS Process Documentation/Guidances (outside of ESOPS)/Others

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.

Coates LC, Helliwell PS2. Defining Low Disease Activity States in Psoriatic Arthritis using Novel Composite Disease Instruments. J Rheumatol. 2016 Feb;43(2):371-5.

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.

Gladman DD, Mease PJ, Ritchlin CT et al, (2007) Adalimumab for Long-Term Treatment of

Psoriasis Arthritis. Arthritis Rheum; 56:476- 88

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

Hansen IMJ, Emamifar A, Andreasen RA, Antonsen S (2017). No further gain can be achieved by calculating Disease Activity Score in 28 joints with high-sensitivity assay of C-reactive protein because of high intraindividual variability of C-reactive protein, Medicine (2017) 96:1.

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.

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

[redacted]

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)

McInnes IB, Mease PJ, Kirkham B, et al (2015). Secukinumab, a human anti-interleukin-17A

monoclonal antibody, in patients with psoriatic arthritis (FUTURE 2): a randomised, doubleblind, placebo-controlled, phase 3 trial. Lancet; 386: 1137–46

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

Mumtaz A, Gallagher P, Kirby B et al (2011) Development of a preliminary composite disease activity index in psoriatic arthritis. Ann Rheum Dis 2011;70:272–277. doi:10.1136/ard.2010.129379

Program Case Retrieval Sheet available in Cabinets/CREDI Projects/A/AIN457A/Integrated Medical Safety

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

Schoels M, Aletaha D, Funovits J, Kavanaugh A, Baker D,Smolen JS (2010). Application of the DAREA/DAPSA score for assessment of disease activity in psoriatic arthritis, Ann Rheum Dis 2010;69:1441–1447.

Schoels M (2014). Psoriatic arthritis indices. Clin Exp Rheumatol 2014; 32 (Suppl. 85):

S109-S112.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.

Yan X, et al. (2009) Missing data handling methods in medical device clinical trials. Journal of Biopharmaceutical Statistics; 19:6, 1085-1098.

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.