NOVARTIS

Clinical Development

Secukinumab (AIN457)

CAIN457H2315 / NCT02696031

A randomized, double-blind, placebo-controlled multicenter study of secukinumab 150 mg in patients with active non-radiographic axial spondyloarthritis to evaluate the safety, tolerability and efficacy up to 2 years, followed by an optional phase of either 150 mg or 300 mg randomized dose escalation for up to another 2 years

Statistical Analysis Plan (SAP) – Amendment 2

Author: [redacted], Trial Statistician

Document type: SAP Documentation

Document status: Final

Release date: Feb 5, 2019

Number of pages: 81

Document History – Changes compared to previous version of SAP.

Version	Date	Changes
Amendment 1	Jan 23, 2019	-Divided efficacy section 11 for plan A and section 12 for plan B -Changed statistical hierarchy to follow protocol amendment 1 and 2 - Added details on estimand definition for each hypothesis in the testing hierarchy -Added interim analysis details for US testing hierarchy -Added description of group sequential testing for US hierarchy -Added full analysis set 2 for week 52 endpoints interim analysis -Corrected treatment arm description -Changed switch treatment labeling - Added instruction that AEs with onset after switching to standard of care, e.g., TNFα blocker, will still be reported under original treatment according to treatment emergent definition - Modified presentation of cDMARD to follow protocol specification -Changed analysis methodology for week 52 continuous endpoints [redacted] - Added a stepwise approach to handle cases of model convergence issues for logistic regression - Added approaches to cases where MI cannot be performed as planned [redacted] -Updated visit windows for MRI and X-ray
Amendment 2	Feb 5, 2019	[redacted] -Corrected typos and added minor clarifications

Table of contents

List of abbreviations

ACR	American College of Rheumatology
AE	Adverse Event
ALT	Alanine Aminotransferase/ Glutamic Pyruvic Transaminase/ GPT
AS	Ankylosing Spondylitis
ASAS	Assessment of Spondyloarthritis International Society criteria
ASDAS	Ankylosing Spondylitis Disease Activity Score
AST	Aspartate Aminotransferase/ Glutamic Oxaloacetic Transaminase/ GOT
ASQoL	Ankylosing Spondylitis Quality of Life Questionnaire
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BSL	Baseline
[redacted]	[redacted]
cDMARD	conventional (i.e., non-biologic) DMARD
CPO	Country Pharma Organization
CRP (hsCRP)	(high sensitivity) C-Reactive Protein
CRP+	hsCRP > ULN (as defined by the central lab)
DMARD	Disease Modifying Antirheumatic Drug
DMC	Data Monitoring Committee
ECG	Electrocardiogram
eCRF	electronic Case Report/Record Form
EMA / EMEA	European Medicines (Evaluation) Agency
[redacted]	[redacted]
[redacted]	[redacted]
FAS	Full Analysis Set
GCP	Good Clinical Practice
GGT	Gamma Glutamyl Transferase
HLA	Human Leukocyte Antigen
ICH	International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use
IEC	Independent Ethics Committee

IL	Interleukin
IRB	Institutional Review Board
IRT	Interactive Response Technology
MAR	Missing At Random
[redacted]	[redacted]
MCS	Mental Component Summary
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MMRM	Mixed Model Repeated Measures
[redacted]	[redacted]
MRI	Magnetic Resonance Imaging
MRI+	Positive for sacroiliitis (active inflammatory lesions)
MTX	Methotrexate
nr-axSpA	Non-radiographic axial SpondyloArthritis
NRI	Non-Responder Imputation
NSAID	Non-Steroidal Anti-Inflammatory Drug
PCS	Physical Component Summary
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
PoC	Proof of Concept
PRN	According to need, as required
PRO	Patient Reported Outcome
PsA	Psoriatic Arthritis
QoL	Quality of Life
RA	Rheumatoid Arthritis
SAE	Serious Adverse Event
s.c.	Subcutaneous(ly)
SCR	Screening
SF-36	Medical Outcome Short Form (36) Health Survey
SpA	Spondyloarthritis
TNF	Tumor Necrosis Factor

TNF-IR	TNFα Inhibitor Incomplete Responders
ULN	Upper Limit of Normal
VAS	Visual Analog Scale
[redacted]	[redacted]

1 Introduction

Data will be analyzed by Novartis according to the data analysis section 9 of the 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. Separate statistical analysis plan will be prepared for the extension phase analysis.

This document comprises two separate endpoint testing strategies based on regional regulatory precedent and feedback:

- Analysis plan A: EU and other non-USA Regions
- Analysis plan B: USA

2 Study objectives

There are two primary objectives based on regional regulatory precedent and feedback. These objectives will be tested in separate analysis plans.

Analysis plan A

The primary objective is to demonstrate that secukinumab 150 mg (with load) at Week 16 is superior to placebo in TNF naïve patients with active non-radiographic axial spondyloarthritis (nr-axSpA) based on the proportion of patients achieving an ASAS40 (Assessment of Spondyloarthritis International Society criteria) response.

The following are the secondary objectives:

- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of patients meeting the ASAS40 response criteria
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of patients meeting the ASAS 5/6 response criteria
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of patients achieving BASDAI 50
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline of high sensitivity C-Reactive Protein (hsCRP)

- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Functional Index (BASFI)
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from screening in SI joint edema on MRI
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of subjects achieving an ASAS20 response
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in Short Form-36 Physical Component Summary (SF-36 PCS)
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in ASQoL
- To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of patients achieving ASAS partial remission
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of TNF naïve patients meeting the ASAS40 response criteria
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of patients meeting the ASAS40 response criteria
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of patients meeting the ASAS 5/6 response criteria
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of patients achieving BASDAI 50
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline of high sensitivity C-Reactive Protein (hsCRP)
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Functional Index (BASFI)
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from screening in SI joint edema on MRI
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of subjects achieving an ASAS20 response
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in Short Form-36 Physical Component Summary (SF-36 PCS)

- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in ASQoL
- To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of patients achieving ASAS partial remission

Analysis plan B

The primary objective is to demonstrate that secukinumab 150 mg without load at Week 52 is superior to placebo in TNF naïve patients with active non-radiographic axial spondyloarthritis (nr-axSpA) based on the proportion of patients achieving an ASAS40 (Assessment of Spondyloarthritis International Society criteria) response.

The following are the secondary objectives:

1. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 52 is superior to placebo based on the proportion of subjects achieving an ASAS40 response
2. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of subjects achieving an ASAS40 response
3. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 52 is superior to placebo based on the proportion of TNF naïve subjects achieving an ASAS40 response
4. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 52 is superior to placebo based on the proportion of subjects achieving an ASAS40 response
5. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of subjects achieving an ASAS40 response
6. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
7. To demonstrate that the efficacy of secukinumab 150 mg (without load), at Week 16 is superior to placebo based on the proportion of patients achieving BASDAI 50
8. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 52 is superior to placebo based on the proportion of patients achieving BASDAI 50
9. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline of high sensitivity C-Reactive Protein (hsCRP)
10. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in Short Form-36 Physical Component Summary (SF-36 PCS)
11. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in ASQoL
12. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of patients meeting the ASAS 5/6 response criteria
13. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the proportion of subjects achieving an ASAS20 response

14. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Functional Index (BASFI)
15. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 16 is superior to placebo based on the change from screening in SI joint edema on MRI
16. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 52 is superior to placebo based on the proportion of patients achieving Ankylosing Spondylitis Disease Activity Score (ASDAS)-C-Reactive Protein (CRP) inactive disease as defined by ASDAS < 1.3
17. To demonstrate that the efficacy of secukinumab 150 mg s.c., (without load), at Week 52 is superior to placebo based on the change from baseline in Ankylosing Spondylitis Quality of Life (ASQoL) scores
18. To demonstrate that the efficacy of secukinumab 150 mg (without load) at Week 52 is superior to placebo based on the change from screening in SI joint edema on MRI
19. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
20. To demonstrate that the efficacy of secukinumab 150 mg (with load), at Week 16 is superior to placebo based on the proportion of patients achieving BASDAI 50
21. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 52 is superior to placebo based on the proportion of patients achieving BASDAI 50
22. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline of high sensitivity C-Reactive Protein (hsCRP)
23. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in Short Form-36 Physical Component Summary (SF-36 PCS)
24. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in ASQoL
25. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of patients meeting the ASAS 5/6 response criteria
26. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the proportion of subjects achieving an ASAS20 response
27. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Functional Index (BASFI)
28. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 16 is superior to placebo based on the change from screening in SI joint edema on MRI
29. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 52 is superior to placebo based on the proportion of patients achieving Ankylosing Spondylitis Disease Activity Score (ASDAS)-C-Reactive Protein (CRP) inactive disease as defined by ASDAS < 1.3

30. To demonstrate that the efficacy of secukinumab 150 mg s.c.,with load, at Week 52 is superior to placebo based on the change from baseline in Ankylosing Spondylitis Quality of Life (ASQoL) scores
31. To demonstrate that the efficacy of secukinumab 150 mg (with load) at Week 52 is superior to placebo based on the change from screening in SI joint edema on MRI

3 Data presentation

Summary statistics for continuous variables will include N, mean, standard deviation, minimum, lower quartile, median, upper quartile, maximum. Summary statistics for discrete variables will be presented in the number and percent of subjects in each category.

If not otherwise specified, p-values will be presented as two-sided p-values and two-sided confidence intervals will be displayed.

The default level of significance will be set to 5% (two-sided, family-wise type-I-error).

Data analyses will be presented by treatment regimen. Efficacy and safety data for the placebo-controlled period (or the entire treatment period as appropriate) will be presented by the following 3 treatment groups. Subjects may be included in more than one treatment group for some analyses (e.g., exposure-adjusted adverse events over the entire treatment period). These treatment groups represent the regimens subjects will be randomized to for the first 52 weeks of the study.

- Secukinumab regimen 1 (secukinumab 150 mg Load): secukinumab 150 mg (1 mL, 150 mg/mL) s.c. PFS at BSL, Weeks 1, 2, and 3, followed by administration every four weeks starting at Week 4
- Secukinumab regimen 2 (secukinumab 150 mg No Load): secukinumab 150 mg (1 mL, 150 mg/mL) s.c. PFS at BSL, placebo at Weeks 1, 2 and 3, followed by secukinumab 150 mg PFS administration every four weeks starting at Week 4
- Placebo regimen: placebo (1 mL) s.c. PFS at BSL, Weeks 1, 2, 3, followed by administration every four weeks starting at Week 4

Note that the treatment groups above for a subject may differ depending on the time period of the analysis and whether one assesses the subject for efficacy or safety.

Furthermore, additional aspects of efficacy, safety and tolerability of secukinumab will be investigated.

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 CAIN457H2315, this is the first 52-weeks of treatment. Comparative efficacy will be performed based on the FAS population using the randomized treatment. Also, the active secukinumab regimens will be compared using confidence intervals on the FAS population.

Efficacy data following treatment switch

Data will also be presented after treatment switch, by a combination of the 'original' and 'switch' treatment groups. This will represent the treatment combinations the subjects experience over the course of the entire trial in case of switch.

All listings will be presented by randomized treatment but also indicating the current treatment.

4 Subjects and treatments

4.1 Analysis sets

The following analysis sets will be used in this trial:

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

Mis-randomized subjects are defined as those subjects who were mistakenly randomized into the IRT 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 evaluated according to the treatment assigned to at randomization, but actual stratum, if stratified randomization is used.

Full analysis set 2(FAS2): The FAS2 will be comprised of all subjects from the randomized set to whom study treatment has been assigned and who have been in enrolled at least 379(upper limit of visit window for primary endpoint) days before date cut off. Following the intent-to-treat principle, subjects will be evaluated according to the treatment assigned to at randomization, but actual stratum, if stratified randomization is used.

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.

The FAS2 set will be used for analyzing the week 52 efficacy endpoints at interim analysis.

4.2 Treatment groups

The summaries by treatment will primarily be performed by the randomized treatment but also by combination of randomized treatment and switch treatment. For some safety summaries (e.g., exposure-adjusted) the 'switch' treatment may be summarized separately.

- Randomized treatment:
  - AIN457 150 mg Load
  - AIN457 150 mg No Load
  - Placebo
- Treatment sequence:
  - AIN457 150 mg Load

  - AIN457 150 mg No Load
  - Placebo – AIN457 150 mg No Load
  - Placebo – Standard of care
  - AIN457 150 mg Load – Standard of care
  - AIN457 150 mg No Load – Standard of care
- Switch treatments (for patients who cross-over):
  - Open label AIN457 150 mg
  - Standard of care

5 Subgroup definitions

The primary endpoint(s) and secondary endpoints will be evaluated within TNFα inhibitor status (TNF-naïve and TNF-IR) and within stratification factor levels (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+).

6 Assessment windows, baseline and post baseline definitions, missing data handling

Baseline and post-baseline definitions

In general, a *baseline* value refers to the last measurement made prior to administration of the first dose of study treatment. A post-baseline value refers to a measurement taken after the first dose of study treatment.

Analysis visit windows

Analysis visit windows will be used for the data that is summarized by visit; they are based on the study evaluation schedule and comprise a set of days around the nominal visit day. For any assessment, there are protocol defined scheduled visits around which analysis visit windows were created to cover the complete range of days within the study. The analysis visit windows and rules for dealing with multiple measurements within the windows are described in Appendix 17.1.

7 Subject disposition, background and demographic characteristics

7.1 Subject 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 (Week 24, Week 52, Week 104), if appropriate, 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.

7.2 Background and demographic characteristics

The following common background and demographic variables will be summarized:

Continuous variables:

- Age (which is derived from date of birth and the date of informed consent)
- Height
- Weight
- Body mass index (BMI) = (body weight in kilograms) / (height in meters)^2

For BMI, height and body weight, the last value prior to randomization is used. If there is no weight recorded prior to taking of study drug, BMI will be missing.

Categorical variables:

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

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

- Patient's global assessment of disease activity and other ASAS components, hsCRP (mg/L and >ULN), [redacted], prior use (yes/no) of TNF-alpha inhibitors, use (yes/no) and separate dose of MTX (mg/week), sulfasalazine (g/day) and systemic corticosteroids (mg/day) at randomization, time since first diagnosis of nr-axSpA (years), time since onset of back pain, modified New York criteria for AS, HLA-B27, [redacted] total back pain (VAS), nocturnal back pain (VAS), total BASDAI score, spinal pain (BASDAI question #2) and [redacted] presence of SIJ inflammation by MRI and each randomization strata level (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+).

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.

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

8 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 nr-axSpA medical history will also be provided.

Smoking history will be summarized by treatment group.

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

9 Study treatment

The analysis of study treatment data will be based on the safety set. The number of active and placebo injections will be summarized by treatment group. 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 will be defined as the time from first dose of study 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) will have exposure to both medications 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.

10 Concomitant medication

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 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.

The number and percentage of subjects receiving prior nr-axSpA therapy will be presented by randomized treatment group and, at least for TNFα, also by the reasons for stopping their therapies (primary lack of efficacy, secondary lack of efficacy, lack of tolerability, other). NSAID, glucocorticoid and DMARD (MTX, Leflunomide, Sulfasalazine) use during the study will be summarized.

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

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

Presentation by dose

Methotrexate, sulfasalazine and systemic corticosteroid intake at randomization will be presented based on the amount taken per time unit, e.g., mg/day.

Since different steroids have different strengths the following multiplication factors will be used to convert a dose in mg into a prednisone equivalent dose in mg:

Cortisone (0.20), hydrocortisone (0.25), prednisolone (1.0), triamcinolone (1.25), methylprednisolone (1.25), dexamethasone (6.67), betamethasone (8.33) (corticosteroid converter website).

The reported dose and frequency of intake will be converted into the desired units. If the frequency is missing or specified as 'per needed', 'unknown', 'once', 'other' or if the dose or dose unit is missing, then the medication will not be part of the presentation.

11 Efficacy evaluation for analysis plan A

11.1 Description of efficacy variables

Assessment of Spondyloarthritis International Society criteria (ASAS) response criteria

The ASAS response measures consist of the following assessment domains (Sieper 2009)

Main ASAS domains:

1. Patient's global assessment of disease activity measured on a VAS scale
2. Patient's assessment of back pain, represented by either total or nocturnal pain scores, both measured on a VAS scale. For ASAS response analyses, the total back pain will be used.
3. Function represented by BASFI average of 10 questions regarding ability to perform specific tasks as measured by VAS scale
4. Inflammation represented by mean duration and severity of morning stiffness, represented by the average of the last 2 questions on the 6-question BASDAI as measured by VAS scale (at least one question is needed)

Additional assessment domains:

5. Spinal lateral flexion, represented by one of the components of BASMI
6. C-reactive protein (acute phase reactant)

[redacted] ASDAS-CRP, [redacted], clinically important and major improvement

The [redacted] ASDAS-CRP (Ankylosing Spondylitis Disease Activity Score) will be utilized to assess the disease activity status. Parameters used for the ASDAS include: total back pain (BASDAI question 2), the patient global assessment of disease activity (ASAS component 1), peripheral pain/swelling (BASDAI question 3), duration of morning stiffness (BASDAI question 6) and high sensitivity C-Reactive Protein (hsCRP) [redacted] (Sieper 2009, Lukas 2009).

The ASDAS formulas are as follows:

ASDAS-CRP = 0.121 x total back pain + 0.110 x patient global + 0.073 x peripheral pain/swelling + 0.058 x duration of morning stiffness + 0.579 x ln(hsCP +1)

[redacted]

hsCRP is in mg/liter, [redacted] the range of other variables is from 0 to 10; ln represents the natural logarithm; √ represents the square root. If any of the ASDAS components are missing ASDAS will not be calculated.

Disease activity states: inactive disease, moderate disease activity, high disease activity, and very high disease activity. The 3 values selected to separate these states were < 1.3 between inactive disease and moderate disease activity, < 2.1 between moderate disease activity and high disease activity, and > 3.5 between high disease activity and very high disease activity. Selected cutoffs for improvement scores were a change ≥ 1.1 unit for "minimal clinically important improvement" and a change ≥ 2.0 units for "major improvement" (Machado 2011). Separate states will be calculated for ASDAS-CRP [redacted].

ASAS components

Patient's global assessment of disease activity (VAS)

The patient's global assessment of disease activity will be performed using a 100 mm visual analog scale (VAS) ranging from not severe to very severe, after the question *"How active was your disease on average during the last week?*".

Patient's assessment of total inflammatory back pain and nocturnal back pain intensity (VAS)

The patient's assessment of inflammatory back pain will be performed using a 100 mm VAS ranging from no pain to unbearable pain, after the question "*Based on your assessment, please indicate what is the amount of back pain at any time that you experienced during the last week?*" and *"Based on your assessment, please indicate what is the amount of back pain at night that you experienced during the last week?"*.

Bath Ankylosing Spondylitis Functional Index (BASFI)

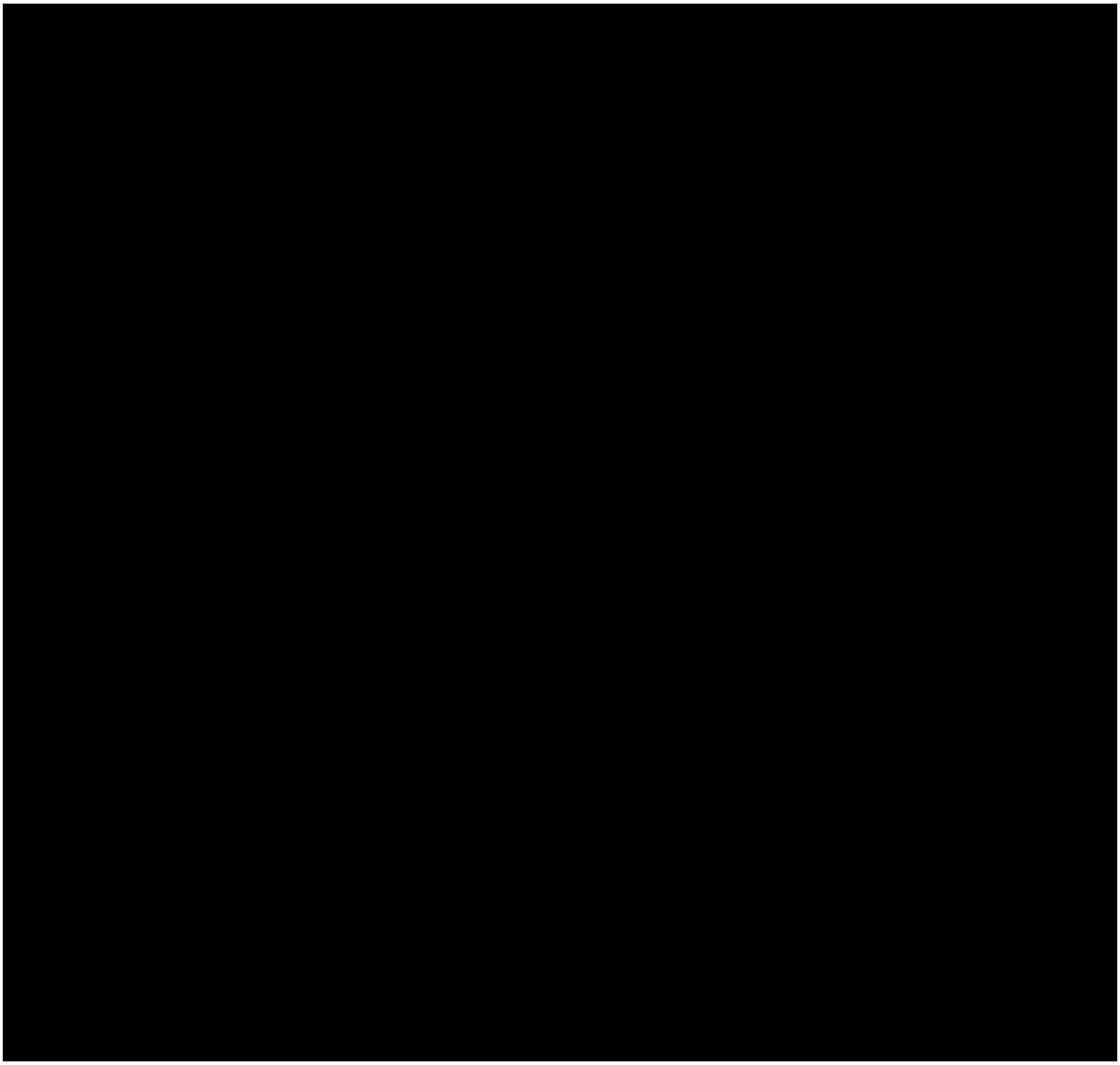
The BASFI is a set of 10 questions designed to determine the degree of functional limitation in those subjects with AS. The ten questions were chosen with a major input from subjects with AS. The first 8 questions consider activities related to functional anatomy. The final 2 questions assess the subjects' ability to cope with everyday life. A 100 mm VAS is used to answer the questions. The mean of the ten questions gives the BASFI score – a value between 0 and 10. In the case that some of the BASFI questions are missing, then the average of the non-missing items is used (Braun 2009, van Tubergen 2001).

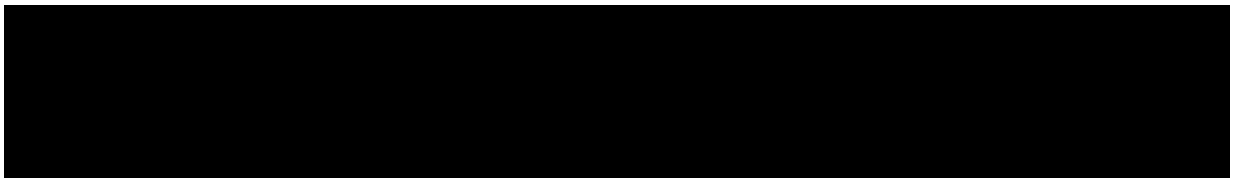
Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)

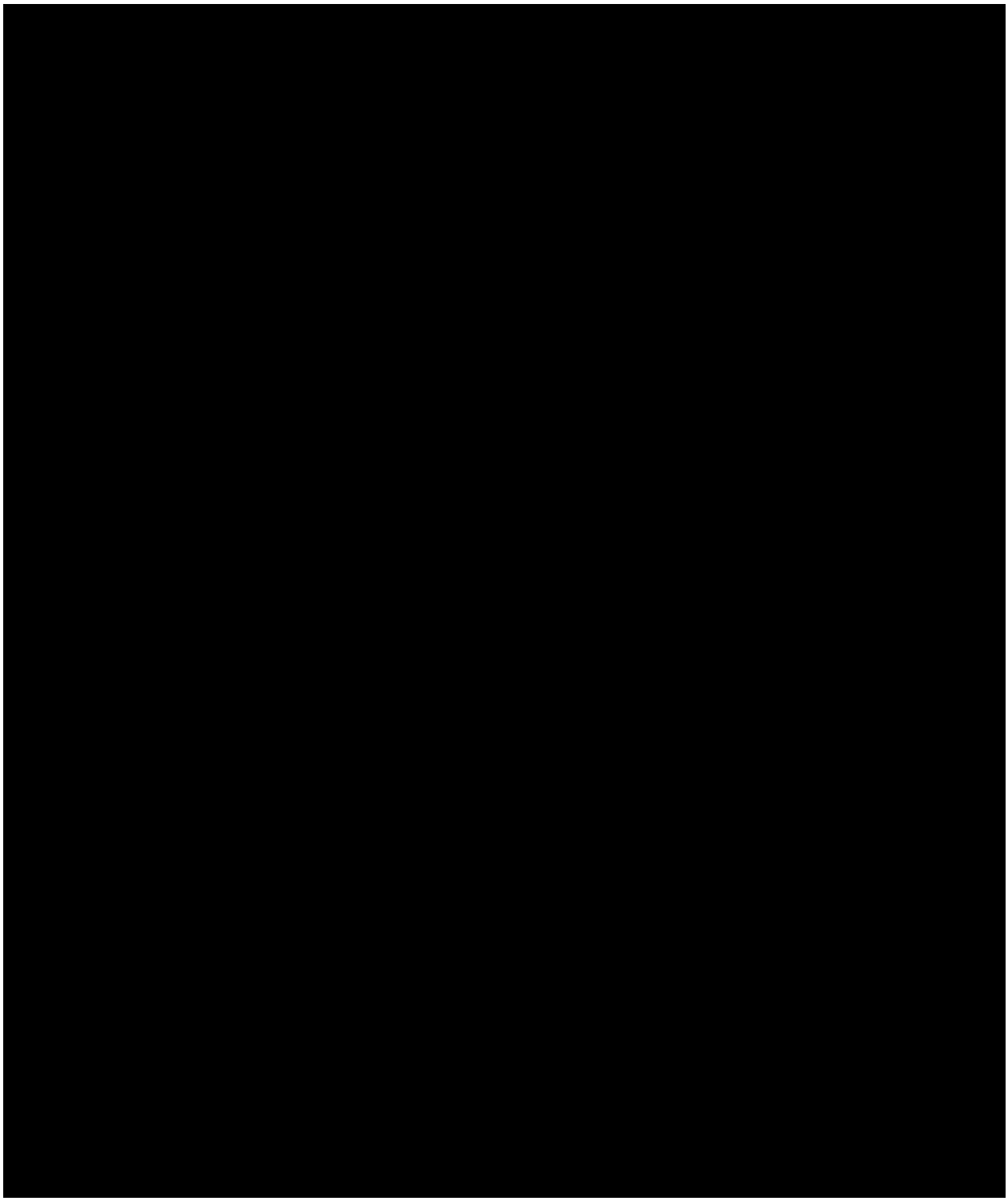
The gold standard for measuring and evaluating disease activity in AS is the BASDAI. The BASDAI consists of a 0 through 10 scale (0 being no problem and 10 being the worst problem), which is used to answer 6 questions pertaining to the 5 major symptoms of AS:

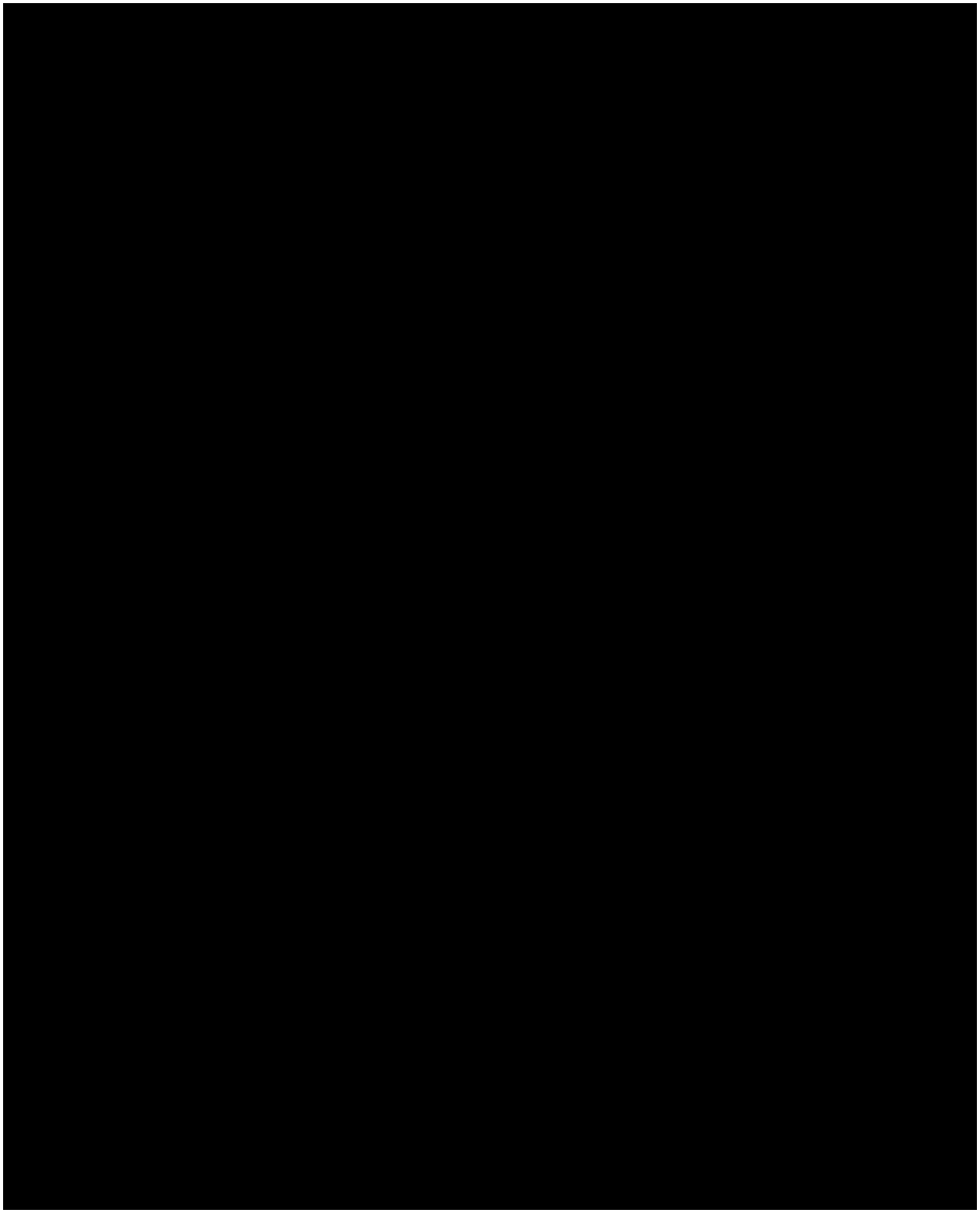
1. Fatigue
2. Spinal pain
3. Joint pain / swelling
4. Areas of localized tenderness
5. Morning stiffness severity
6. Morning stiffness duration

To give each symptom equal weighting, the mean (average) of the two scores relating to morning stiffness is taken and is then added to the sum of the first 4 questions. The resulting 0 to 50 score is divided by 5 to give a final 0 – 10 BASDAI score. Scores of 4 or greater suggest suboptimal control of disease, and subjects with scores of 4 or greater are usually good candidates for either a change in their medical therapy or for enrollment in clinical trials evaluating new drug therapies directed at AS. BASDAI is a quick and simple index taking between 30 secs and 2 mins to complete. At least 4 questions should be non-missing to calculate the BASDAI score. Otherwise, BASDAI score will be missing (Haywood 2002). If both Q5 and Q6 are missing or one of Q1 to Q4 is missing, then the total sum should be divided by 4 instead of 5. If two of Q1 to Q4 is missing and both Q5 and Q6 are not missing, then the sum should be divided by 3.









11.2 Description of health-related quality of life variables

Ankylosing Spondylitis Quality of Life (ASQoL)

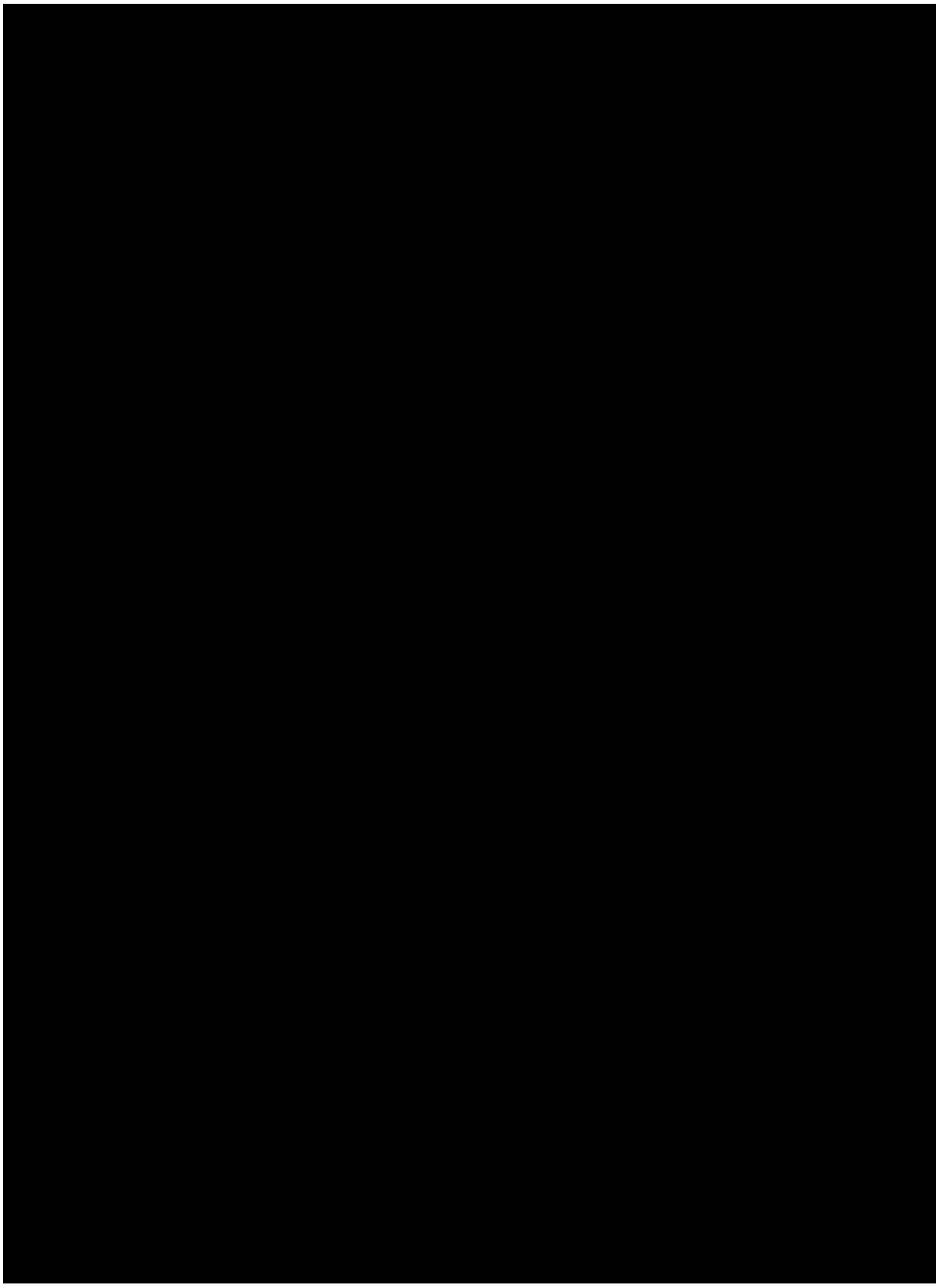
The ASQoL is a self-administered questionnaire designed to assess health-related quality of life in adult patients with Ankylosing Spondylitis. The ASQoL contains 18 items with a dichotomous yes/no response option. A single point is assigned for each "yes" response and no points for each "no" response resulting in overall scores that range from 0 (least severity) to 18 (highest severity). As such, lower score indicate better quality of life. Items include an assessment of mobility/energy, self-care and mood/emotion. The recall period is "at the moment," and the measure requires approximately 6 minutes to complete.

At least 15 answered questions are required to be able to calculate ASQoL using mean imputation, (sum of answered)/(number answered)*18 (Doward 2003)

SF-36

The Short Form Health Survey (SF-36) is a widely used and extensively studied instrument to measure health-related quality of life among healthy subjects and patients with acute and chronic conditions. It consists of eight subscales (domains) that can be scored individually: Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role- Emotional, and Mental Health. Two overall summary scores, the Physical Component Summary (PCS) and the Mental Component Summary (MCS) also can be computed. The SF-36 has proven useful in monitoring general and specific populations, comparing the relative burden of different disease, differentiating the health benefits produced by different treatments, and in screening individual patients. The eight domains are based on a scale from 0-100 while PCS and MCS are norm-based scores with a mean of 50 and a standard deviation of 10.

Quality metric uses weighted maximum likelihood estimation, a modified version of item response theory (IRT) to estimate scale scores when a respondent is missing multiple items. The PCS summary score measure requires scores for seven scales, one of which must be the PF scale and the MCS score also requires scores for seven scales, one of which must be the MH scale. Only one item is needed for each of the multi-item domains.



11.3 Handling of missing data

Missing data for ASAS20/40 response and other binary efficacy variables (e.g., ASAS5/6, etc.) for data up to Week 20 will be handled as follows:

1. Patients who drop out of the trial for any reason will be considered as non-responders from the time they drop out through Week 20
2. Patients who do not have the required data to compute responses (e.g., ASAS components) at baseline and at the specific timepoint will be classified as non-responders at the specific timepoint.

Patients who are unblinded will be considered non-responders from the time of unblinding up to Week 20. The primary analysis will use non-responder imputation.

Continuous variables (e.g., ASAS components) with the exception of MRI endpoints, will be analyzed using a mixed-effects model repeated measures (MMRM) which is valid under the missing at random (MAR) assumption. The MMRM models will be applied up to week 20 when no treatment escape has occured. 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 patient will be removed from the analysis of the corresponding variable, i.e., it might be that the number of patients providing data to an analysis is smaller than the number of patients in the FAS.

For SI joint edema on MRI a multiple imputation (MI) approach under MAR assumption will be applied to handle missing data. The MI model will include stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α inhibitor status as categorical covariates and patient weight as a continuous covariate.

Imputation under MAR relies on the assumption that unbiased estimates can be obtained by borrowing information from patients with collected data that are similar in regards to model baseline covariates and measurements collected at prior visits.

11.4 Analysis of primary variable

Description

The ASAS Response Criteria (ASAS40) is defined as an improvement of ≥40% and ≥2 unit on a scale of 10 in at least three of the four main domains and no worsening at all in the remaining domain.

Primary endpoint for analysis plan A

The estimand for primary endpoint is defined as follows

A. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted nr-axSpA population

B. Variable: composite of remaining on the study and on randomized treatment through 16 weeks and achieving ASAS40 response at 16 weeks

C. Intercurrent event: the intercurrent event is captured through the variable definition

D. Population-level summary: odds ratio of response proportions between the treatment conditions

The statistical hypothesis for ASAS40 being tested is that there is no difference in the proportion of TNF naïve patients fulfilling the ASAS40 criteria at Week 16 in the secukinumab 150 mg load regimen versus placebo regimen.

Let p_j denote the proportion of ASAS40 responders at Week 16 in TNF naïve patients for treatment regimens j, j=0, 1 where

- 0 corresponds to placebo regimen,
- 1 corresponds to secukinumab 150 mg with loading regimen,

In statistical terms, H_1: $p_1 = p_0$, H_{A1}: $p_1 \neq p_0$, i.e.,

H_1: secukinumab 150 mg with load is not different to placebo regimen with respect to ASAS40 response in TNF naïve patients at Week 16

The primary analysis will be conducted via logistic regression with treatment and stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) as factors and weight as a covariate. Odds ratios and 95% CI will be presented comparing each secukinumab regimen to placebo.

Odds ratio and 95% CI will also be presented comparing the two secukinumab regimens.

11.5 Analysis of secondary variables

Secondary endpoints for analysis plan A

The secondary efficacy variables include,

- response to treatment at Week 16 according to the ASAS40 criteria
- response to treatment at Week 16 according to the ASAS 5/6 criteria
- change from baseline in total BASDAI score at Week 16
- response to treatment at Week 16 according to the BASDAI50 criteria
- change from baseline in hsCRP at Week 16
- change from baseline in total BASFI score at Week 16
- change from screening in MRI SI joint edema score at Week 16
- response to treatment at Week 16 according to the ASAS20 criteria
- change from baseline in SF-36 PCS score at Week 16
- change from baseline in ASQoL score at Week 16
- response to treatment at Week 16 according to the ASAS partial remission

All analysis will be done in the FAS population.

Testing strategy to control type I error

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

Primary objective:

H_1: secukinumab 150 mg (with load) is not different to placebo regimen with respect to signs and symptoms (ASAS40 response) in TNF naïve patients at Week 16

Secondary objectives:

H_2: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS40 response at Week 16

H_3: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS 5/6 response at Week 16

H_4: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in total BASDAI at Week 16

H_5: secukinumab 150 mg (with load) is not different to placebo regimen with respect to BASDAI50 at Week 16

H_6: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in hsCRP at Week 16

H_7: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in BASFI at Week 16

H_8: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from screening in SI joint edema on MRI at Week 16

H_9: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS20 at Week 16

H_{10}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in SF-36 PCS at Week 16

H_{11}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in ASQoL at Week 16

H_{12}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS partial remission at Week 16

H_{13}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS40 response in TNF naïve patients at Week 16

H_{14}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS40 response at Week 16

H_{15}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS 5/6 response at Week 16

H_{16}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in total BASDAI at Week 16

H_{17}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to BASDAI50 at Week 16

H_{18}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in hsCRP at Week 16

H_{19}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in BASFI at Week 16

H_{20}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from screening in SI joint edema on MRI at Week 16

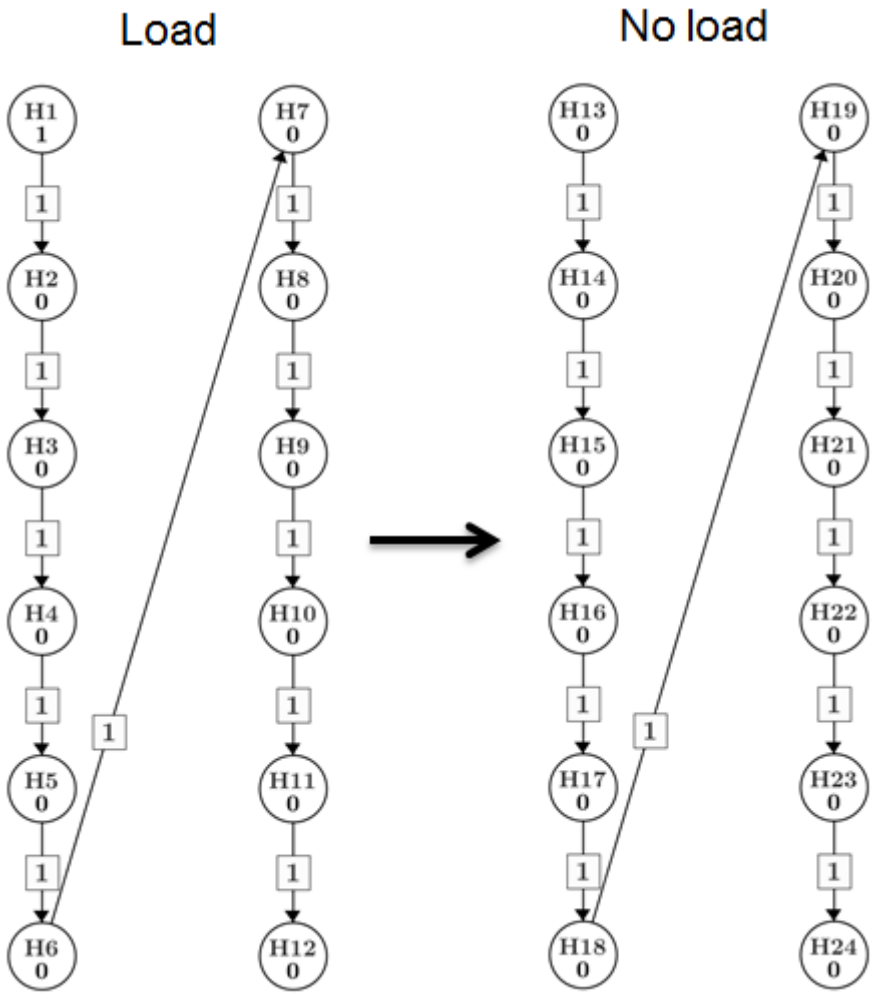
H_{21}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS20 at Week 16

H_{22}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in SF-36 PCS at Week 16

H_{23}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in ASQoL at Week 16

H_{24}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS partial remission at Week 16

Figure 11-1 Testing strategy for analysis plan A



The family-wise error will be set to α=5% and it will be controlled with the proposed sequential testing strategy as described in Figure 11-1. The primary hypothesis (H_1) for the primary objective (ASAS40 in TNF naïve at Week 16) for secukinumab with load regimen versus placebo will be tested at α-level. If the hypothesis H_1 is rejected then the whole α will be passed to the next hypothesis (H_2) which will be tested at α-level. This procedure will continue (pending rejection of the null hypotheses) until H_{12} is rejected. If H_{12} is rejected then the full α-level is passed on to the testing sequence of secukinumab without load which can now be tested at 5% level sequentially in a similar way.

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 sequence for the test of another hypothesis if the treatment effect is in favor of secukinumab.

The secondary efficacy variables and the method of analyses are described below.

Estimand definition for the secondary variables

Estimand definition for the secondary binomial variables is the following

A. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted nr-axSpA population

B. Variable: composite of remaining on the study and on randomized treatment through 16 weeks and achieving *variable* response at 16 weeks

C. Intercurrent event: the intercurrent event is captured through the variable definition

D. Population-level summary: odds ratio of response proportions between the treatment conditions

The estimand of binary variables is (each secukinumab dose vs placebo) obtained from a logistic regression model at week 16 in the FAS population. In the analysis patients dropping out or being unblinded before week 16 or having missing response data at week 16 are considered as non-responders.

Estimand definition for the secondary continuous variables is the following

A. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted nr-axSpA population

B. Variable: change from baseline in the *variable* of interest

C. Intercurrent event: had no intercurrent events occured before week 16

D. Population-level summary: difference in variable means between the treatment conditions

The estimand of continuous variables at week 16 is (each secukinumab dose vs placebo) obtained from a linear regression model in the FAS population assuming patients dropping out or having missing data at week 16 are missing-at-random (MAR). MAR imputation will be implemented through an MMRM or multiple imputation model.

ASAS40 at Week 16

The proportion of patients meeting the response criteria will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and weight as a covariate.

ASAS 5/6 at Week 16

The proportion of patients meeting the response criteria at Week 16 will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and weight as a covariate.

BASDAI at Week 16

Between-treatment differences in the change from baseline in BASDAI will be evaluated using MMRM with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+),

TNF-α status, and analysis visit as factors and baseline BASDAI score and weight as continuous covariates. Treatment by analysis visit and baseline BASDAI score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. The significance of the treatment effect for secukinumab regimens will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at Week 16.

BASDAI50 at Week 16

The proportion of patients meeting the response criteria at Week 16 will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and baseline weight and BASDAI score as covariates.

hsCRP at Week 16

For the change in hsCRP, since evidence from the literature would suggest that the data are not normally distributed (Huffman et al 2006), analysis will be performed on the $\log_e$ ratio of the treatment value vs baseline value (calculated by dividing the post-baseline value by the baseline value and then applying the $\log_e$ transformation) to normalize the distribution of hsCRP at each analysis visit. Between-treatment differences in the change in hsCRP relative to baseline will be evaluated using MMRM with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status, and analysis visit as factors and $\log_e$ baseline hsCRP and weight as continuous covariates. Treatment by analysis visit and $\log_e$ baseline hsCRP by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. The significance of the secukinumab treatment effect for secukinumab regimens will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at Week 16. The estimate and the 2-sided 95% confidence intervals obtained from the model will be back transformed to the original scale.

BASFI at Week 16

Between-treatment differences in the change from baseline in BASFI will be evaluated using MMRM with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status, and analysis visit as factors and baseline BASFI score and weight as continuous covariates. Treatment by analysis visit and baseline BASFI score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. The significance of the treatment effect for secukinumab regimens will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at Week 16.

SI joint edema on MRI at Week 16

The change from baseline to Week 16 in inflammation measured by SI joint total edema score will be evaluated using an ANCOVA model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors, and weight and baseline inflammation score as covariates.

[redacted]

[redacted]

ASAS20 at Week 16

The proportion of patients meeting the response criteria at Week 16 will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and weight as a covariate.

SF-36 PCS at Week 16

See Section 11.7.

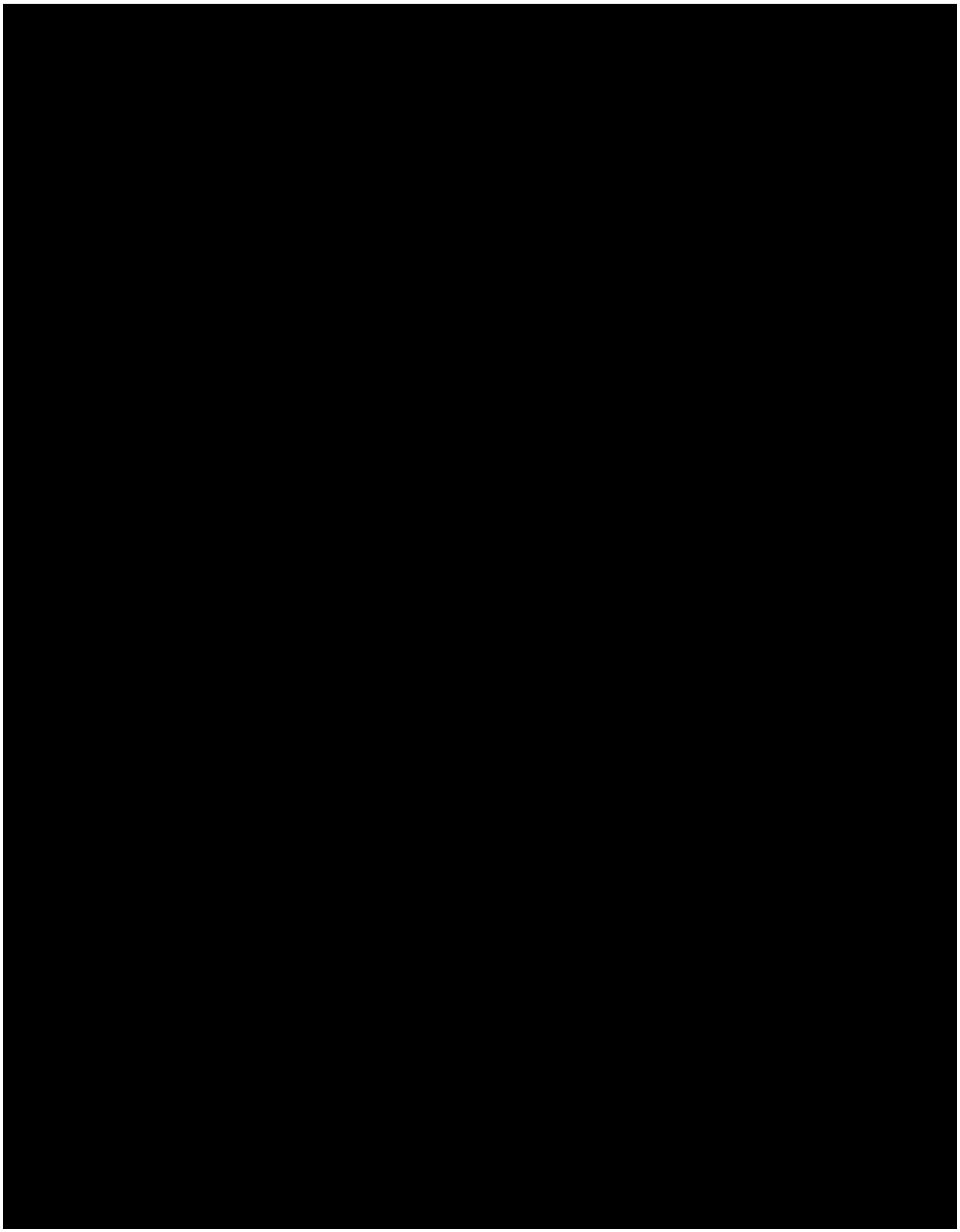
ASQoL at Week 16

See Section 11.7.

ASAS partial remission at Week 16

Response at Week 16 to ASAS partial remission will be evaluated using a logistic regression model with treatment, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and weight as a covariate.

[redacted]



11.7 Analysis of health-related quality of life endpoints

Ankylosing Spondylitis Quality of Life (ASQoL)

ASQoL at week 16

For the change in ASQoL scores, between-treatment differences in the change in ASQoL scores will be evaluated using MMRM. Treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status and analysis visit will be used as categorical factors and baseline ASQoL score and weight as continuous covariates. Treatment by analysis visit and baseline ASQoL score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the 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 and/or secukinumab at the appropriate analysis visits.

SF-36

The following variables will be evaluated:

- SF-36 domain scores (based on a scale of 0-100).
- SF-36 PCS and MCS scores (norm-based scores).
- SF-36 PCS responder (improvement of ≥ 2.5 points, Lubeck 2004)

For the change in SF-36 summary scores (PCS and MCS), between-treatment differences will be evaluated using MMRM. Treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status and analysis visit will be included as categorical factors and baseline SF-36 score (PCS or MCS) and weight as continuous covariates. Treatment by analysis visit and baseline SF-36 score (PCS or MCS) by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. Pairwise comparisons will be performed between secukinumab regimens and placebo and/or secukinumab at appropriate analysis visits.

In the responder analyses, treatment groups will be compared with respect to response to treatment using a logistic regression model with treatment, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status as factors and baseline SF-36 score (PCS or MCS) and weight as covariates. Odds ratios and 95% CI will be presented for appropriate treatment comparisons.

The SF-36 domain scores will be summarized by treatment.

12 Efficacy evaluation for analysis plan B

12.1 Description of efficacy variables

Please see section 11.1.

12.2 Description of health-related quality of life variables

Please see section 11.2

12.3 Handling of missing data

Missing data for ASAS20/40 response and other binary efficacy variables (e.g., ASAS5/6, etc.) for data up to Week 20 will be handled as follows:

1. Patients who drop out of the trial for any reason will be considered as non-responders from the time they drop out through Week 52
2. Patients who do not have the required data to compute responses (e.g., ASAS components) at baseline and at the specific timepoint will be classified as non-responders at the specific timepoint.

Patients who are unblinded will be considered non-responders from the time of unblinding up to Week 52. The primary analysis will use non-responder imputation.

Continuous variables (e.g., ASAS components) with the exception of MRI endpoints, will be analyzed using a mixed-effects model repeated measures (MMRM) which is valid under the missing at random (MAR) assumption. The MMRM models will be applied up to week 20 when no treatment escape has occured. 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 patient will be removed from the analysis of the corresponding variable, i.e., it might be that the number of patients providing data to an analysis is smaller than the number of patients in the FAS.

For SI joint edema on MRI a multiple imputation (MI) approach under MAR assumption will be applied to handle missing data. The MI model will include stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α inhibitor status as categorical covariates and patient weight as a continuous covariate.

Imputation under MAR relies on the assumption that unbiased estimates can be obtained by borrowing information from patients with collected data that are similar in regards to model baseline covariates and measurements collected at prior visits

For continuous endpoints analyzed at week 52 with composite strategy missing data will be handled similarly as escape (extreme unfavorable value will be assigned for any missing value).

Data post treatment switch

Within analysis plan B, the statistical testing hierarchy has endpoints at Week 16 and Week 52.

In general, data for patients where a decision was made to escape to biologic treatment (on or after week 20) will be handled in the following fashion up to Week 52 (of note, changes in background medications other than biologics will not be treated as escape):

- For binary endpoints, composite estimand strategy will be applied and patients will be considered non-responders after decision to stop blinded study treatment (this also includes secukinumab patients where switch decision was made although they will stay on

secukinumab). This is referred to as the 'escape penalty'. This will be done for all treatment regimens in order to minimize bias for between-treatment comparisons.

- For continuous endpoints in the testing hierarchy(ASQoL and SI joint edema score on MRI), a composite estimand strategy will be applied and data after switch will be set to the extreme unfavorable value. Rank-based analysis will be used for hypothesis testing.
- Difference in trimmed means estimate will be used to summarize the treatment effect estimate for the continuous composite endpoint
- For binary and continuous endpoints the actual observed values after switch will be also summarized with descriptive statistics

12.4 Analysis of primary variable

Description

The ASAS Response Criteria (ASAS40) is defined as an improvement of ≥40% and ≥2 unit on a scale of 10 in at least three of the four main domains and no worsening at all in the remaining domain.

Primary endpoint for analysis plan B

The primary estimand is defined as follows

A. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted nr-axSpA population

B. Variable: composite of remaining on the study and on randomized treatment through 52 weeks and achieving ASAS40 response at 52 weeks

C. Intercurrent event: the intercurrent event is captured through the variable definition

D. Population-level summary: odds ratio of response proportions between the treatment conditions

The statistical hypothesis for ASAS40 being tested is that there is no difference in the proportion of TNF naïve subjects fulfilling the ASAS40 criteria at Week 52 in secukinumab 150 mg without load regimen versus placebo regimen.

Let p_j denote the proportion of ASAS40 responders at Week 52 for treatment regimens j, j=0, 1, where

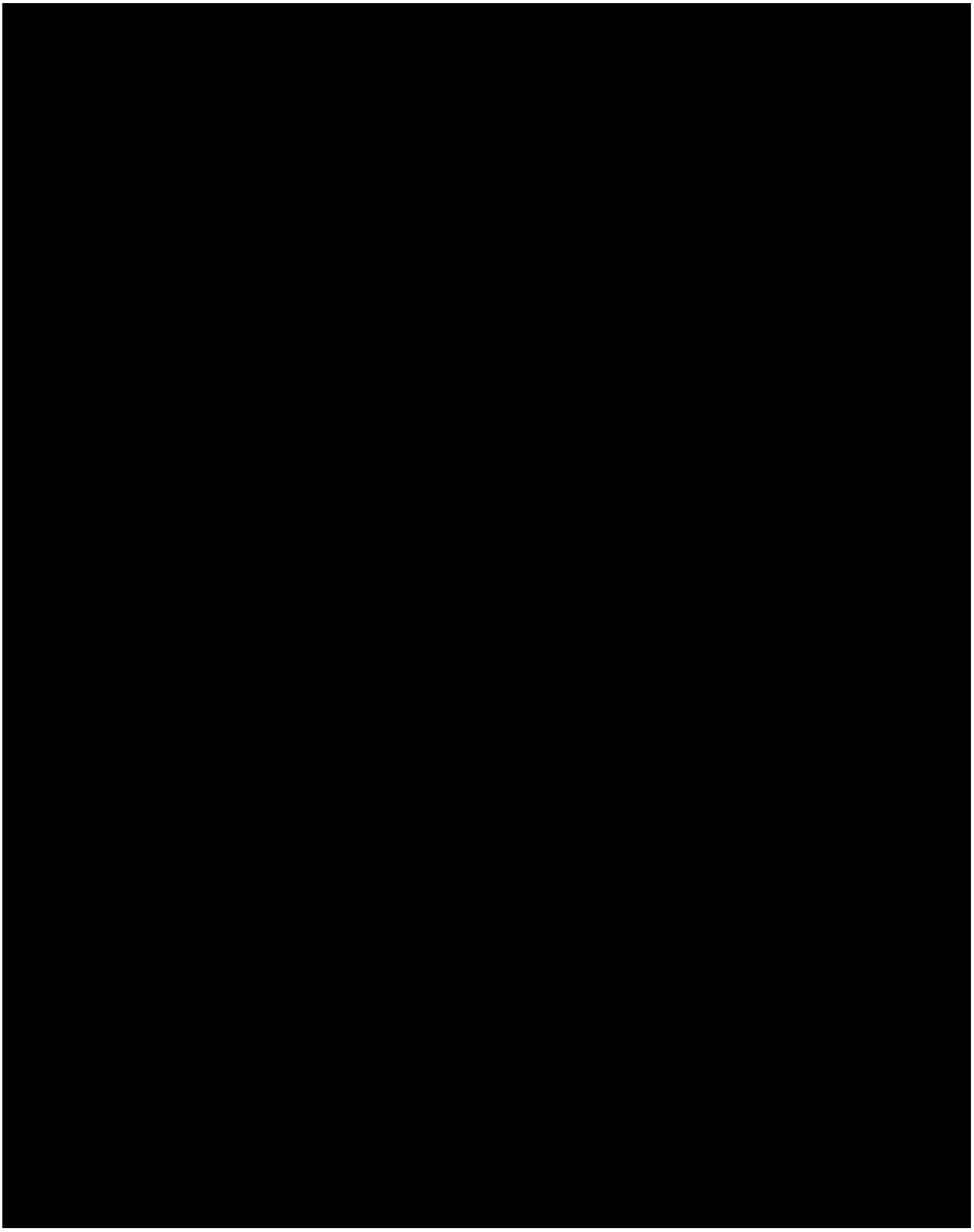
- 0 corresponds to placebo regimen,
- 1 corresponds to secukinumab 150 mg without loading regimen,

In statistical terms, H_1: $p_1 = p_0$, H_{A1}: $p_1 \neq p_0$, i.e.,

H_1: secukinumab 150 mg without load is not different to placebo regimen with respect to ASAS40 response in TNF naïve patients at Week 52.

The primary analysis will be conducted via logistic regression with treatment and stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) as factors and weight as a covariate. Odds ratios and 95% CI will be presented comparing each secukinumab regimen to placebo.

Odds ratio and 95% CI will also be presented comparing the two secukinumab regimens.



12.5 Analysis of secondary variables

Secondary endpoints for analysis plan B

The secondary efficacy variables include,

- response to treatment at Week 52 according to the ASAS40 criteria
- response to treatment at Week 16 according to the ASAS40 criteria
- change from baseline in total BASDAI score at Week 16
- response to treatment at Week 16 according to the BASDAI50 criteria
- response to treatment at Week 52 according to the BASDAI50 criteria
- change from baseline in hsCRP at Week 16
- change from baseline in SF-36 PCS score at Week 16
- change from baseline in ASQoL score at Week 16
- response to treatment at Week 16 according to the ASAS5/6 criteria
- response to treatment at Week 16 according to the ASAS20 criteria
- change from baseline in total BASFI score at Week 16
- change from screening in MRI SI joint edema score at Week 16
- response to treatment at Week 52 according to the ASDAS-CRP Inactive Disease
- change from screening in MRI SI joint edema score at Week 52
- change from baseline in ASQoL score at Week 52

All analysis will be done in the FAS population.

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 objective:

H_1: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS40 in TNF naïve at Week 52

Secondary objectives:

H_2: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS40 at Week 52

H_3: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS40 at Week 16

H_4: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS40 in TNF naïve at Week 52

H_5: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS40 at Week 52

H_6: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS40 at Week 16

H_7: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in total BASDAI at Week 16

H_8: secukinumab 150 mg (without load) is not different to placebo regimen with respect to BASDAI50 at Week 16

H_9: secukinumab 150 mg (without load) is not different to placebo regimen with respect to BASDAI50 at Week 52

H_{10}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in hsCRP at Week 16

H_{11}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in SF-36 PCS at Week 16

H_{12}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in ASQoL at Week 16

H_{13}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS5/6 at Week 16

H_{14}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASAS20 at Week 16

H_{15}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in BASFI at Week 16

H_{16}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from screening in SI joint edema on MRI at Week 16

H_{17}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to ASDAS-CRP ID at Week 52

H_{18}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from baseline in ASQoL at Week 52

H_{19}: secukinumab 150 mg (without load) is not different to placebo regimen with respect to change from screening in SI joint edema on MRI at Week 52

H_{20}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in total BASDAI at Week 16

H_{21}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to BASDAI50 at Week 16

H_{22}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to BASDAI50 at Week 52

H_{23}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in hsCRP at Week 16

H_{24}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in SF-36 PCS at Week 16

H_{25}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in ASQoL at Week 16

H_{26}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS5/6 at Week 16

H_{27}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASAS20 at Week 16

H_{28}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in BASFI at Week 16

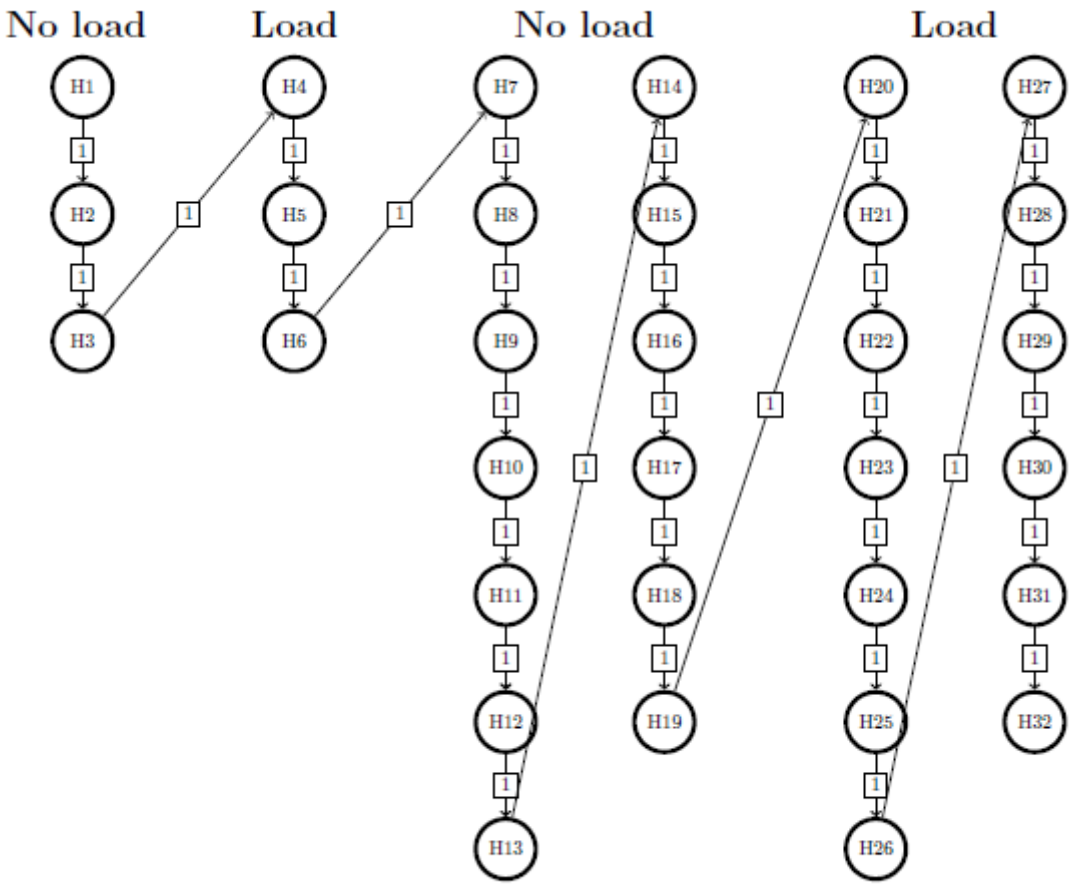
H_{29}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from screening in SI joint edema on MRI at Week 16

H_{30}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to ASDAS-CRP ID at Week 52

H_{31}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from baseline in ASQoL at Week 52

H_{32}: secukinumab 150 mg (with load) is not different to placebo regimen with respect to change from screening in SI joint edema on MRI at Week 52

Figure 12-1 Testing strategy



The family-wise error will be set to α=5% and it will be controlled with the proposed sequential testing strategy as described in Figure 11-2. The primary hypothesis (H_1) for the primary objective (ASAS40 in TNF naïve at Week 52) for secukinumab without load regimen versus placebo will be tested at α-level. If the hypothesis H_1 is rejected then the whole α will be passed to the next hypothesis (H_2) which will be tested at α-level. This procedure will continue (pending rejection of the null hypotheses) until H_{32} is rejected.

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.

Estimand definition

Binomial variables week 16 or week 52

Estimand definition for the secondary binomial variables at 16 (or 52) weeks is the following

A. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted nr-axSpA population

B. Variable: composite of remaining on the study and on randomized treatment through 16 (or 52) weeks and achieving *variable* response at 16 (or 52) weeks

C. Intercurrent event: the intercurrent event is captured through the variable definition

D. Population-level summary: odds ratio of response proportions between the treatment conditions

The estimand of binary variables is (each secukinumab dose vs placebo) obtained from a logistic regression model at week 16 in the FAS population. In the analysis patients dropping out or being unblinded before week 16 (or 52) or having missing response data at week 16 (or 52) are considered as non-responders.

Continuous variables week 16

Estimand definition for the secondary continuous variables at week 16 is the following

A. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted nr-axSpA population

B. Variable: change from baseline in the *variable* of interest

C. Intercurrent event: had no intercurrent events occured before week 16

D. Population-level summary: difference in variable means between the treatment conditions

The estimand of continuous variables at week 16 is (each secukinumab dose vs placebo) obtained from a linear regression model in the FAS population assuming patients dropping out or having missing data at week 16 are missing-at-random (MAR). MAR imputation will be implemented through an MMRM or multiple imputation model

Continuous variables week 52

Estimand definition for the secondary continuous variables at week 52 is the following

A. Population: defined through appropriate inclusion/exclusion criteria to reflect the targeted nr-axSpA population

B. Variable: continuous response variable, for patients experiencing an intercurrent event extreme unfavorable value is assigned

C. Intercurrent event: the intercurrent event is captured through the variable definition;

D. Population-level summary: The difference in ranks supplemented with the trimmed means for change from baseline

Wilcoxon rank-sum test will be used for testing the difference in the composite variable in each secukinumab dose vs placebo. Trimmed means is analyzed using permutation test and MMRM or ANCOVA.

ASAS40 at Week 16 and Week 52

The proportion of patients meeting the response criteria at Week 16 and at Week 52 will be evaluated using a logistic regression model with treatment group, stratification factor

(CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and weight as a covariate.

BASDAI at Week 16

Between-treatment differences in the change from baseline in BASDAI will be evaluated using MMRM with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status, and analysis visit as factors and baseline BASDAI score and weight as continuous covariates. Treatment by analysis visit and baseline BASDAI score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. The significance of the treatment effect for secukinumab regimens will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at Week 16.

BASDAI50 at Week 16 and Week 52

The proportion of patients meeting the response criteria at Week 16 and at Week 52 will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and baseline weight and BASDAI score as covariates.

hsCRP at Week 16

For the change in hsCRP, since evidence from the literature would suggest that the data are not normally distributed (Huffman et al 2006), analysis will be performed on the $\log_e$ ratio of the treatment value vs baseline value (calculated by dividing the post-baseline value by the baseline value and then applying the $\log_e$ transformation) to normalize the distribution of hsCRP at each analysis visit. Between-treatment differences in the change in hsCRP relative to baseline will be evaluated using MMRM with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status, and analysis visit as factors and $\log_e$ baseline hsCRP and weight as continuous covariates. Treatment by analysis visit and $\log_e$ baseline hsCRP by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. The significance of the treatment effect for secukinumab regimens will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at Week 16. The estimate and the 2-sided 95% confidence intervals obtained from the model will be back transformed to the original scale.

SF-36 PCS at Week 16

The following variables will be evaluated:

- SF-36 domain scores (based on a scale of 0-100).
- SF-36 PCS and MCS scores (norm-based scores).
- SF-36 PCS responder (improvement of ≥ 2.5 points, Lubeck 2004)

For the change in SF-36 summary scores (PCS and MCS), between-treatment differences will be evaluated using MMRM. Treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status and analysis visit will be included as categorical factors and baseline SF-36 score (PCS or MCS) and weight as continuous covariates. Treatment by analysis

visit and baseline SF-36 score (PCS or MCS) by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. Pairwise comparisons will be performed between secukinumab regimens and placebo and/or secukinumab at appropriate analysis visits.

In the responder analyses, treatment groups will be compared with respect to response to treatment using a logistic regression model with treatment, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status as factors and baseline SF-36 score (PCS or MCS) and weight as covariates. Odds ratios and 95% CI will be presented for appropriate treatment comparisons.

The SF-36 domain scores will be summarized by treatment.

ASQoL at Week 16

For the change in ASQoL scores, between-treatment differences in the change in ASQoL scores will be evaluated using MMRM. Treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status and analysis visit will be used as categorical factors and baseline ASQoL score and weight as continuous covariates. Treatment by analysis visit and baseline ASQoL score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the 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 and/or secukinumab at the appropriate analysis visits.

ASQoL at Week 52

The change from baseline to Week 52 in ASQoL score will be analyzed by composite estimand strategy in which data after switch is set to extreme unfavorable value. Wilcoxon rank-sum test will be used for testing the difference in distributions of the composite endpoint in active vs placebo. Treatment differences will be summarised using trimmed means and confidence intervals will be calculated using permutation distribution.

[redacted]

ASAS 5/6 at Week 16

The proportion of patients meeting the response criteria at Week 16 will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and weight as a covariate.

ASAS20 at Week 16

The proportion of patients meeting the response criteria at Week 16 will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and weight as a covariate.

BASFI at Week 16

Between-treatment differences in the change from baseline in BASFI will be evaluated using MMRM with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status, and analysis visit as factors and baseline BASFI score and weight as continuous covariates. Treatment by analysis visit and baseline BASFI score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. The significance of the treatment effect for secukinumab regimens will be determined from the pairwise comparisons performed between secukinumab regimens and placebo at Week 16.

SI joint edema on MRI at Week 16

The change from screening to Week 16 in inflammation measured by SI joint total edema score will be evaluated using an ANCOVA model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors, and weight and baseline inflammation score as covariates.

[redacted]

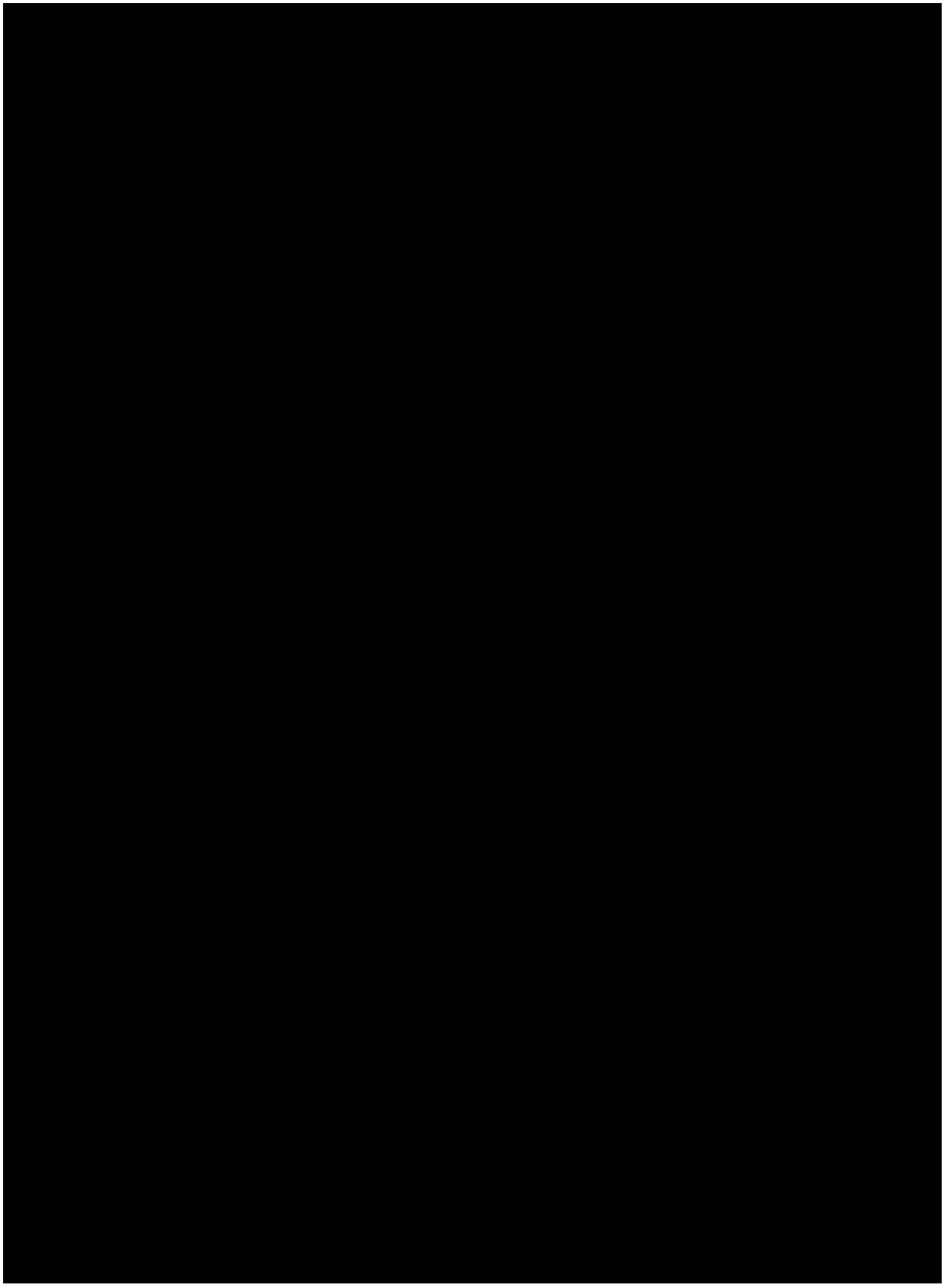
SI joint edema on MRI at Week 52

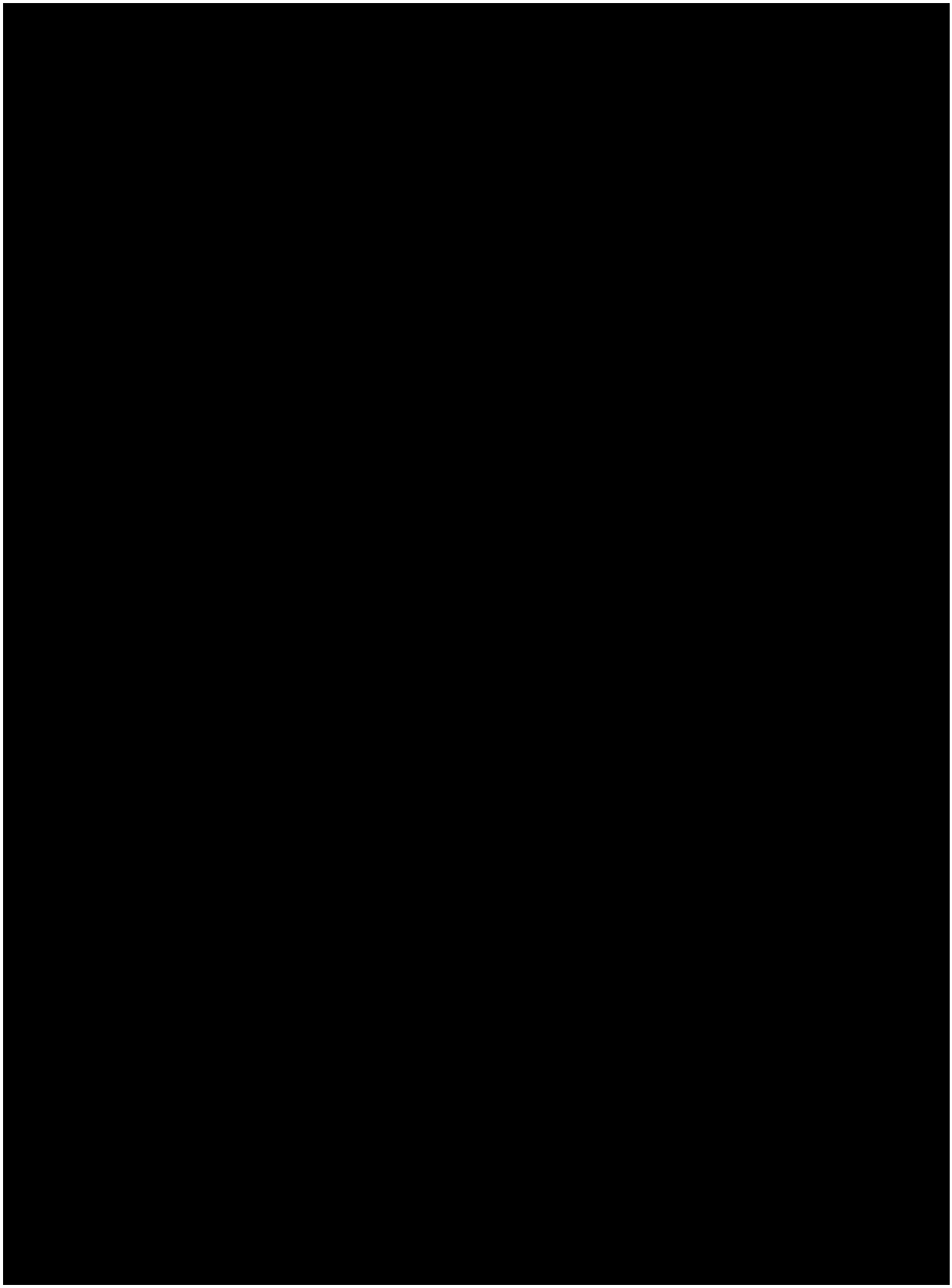
The change from baseline to Week 52 in inflammation measured by SI joint total edema score will be analyzed by composite estimand strategy in which data after switch is set to extreme unfavorable value. Wilcoxon rank-sum test will be used for testing the difference in distributions of the composite endpoint in active vs placebo. Treatment differences will be summarised using trimmed means and confidence intervals will be calculated using permutation distribution.

[redacted]

ASDAS-CRP inactive disease at Week 52

The proportion of patients meeting the response criteria at Week 52 will be evaluated using a logistic regression model with treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+) and TNF-α status as factors and baseline weight and ASDAS-CRP score as covariates.





12.7 Analysis of health-related quality of life endpoints

Ankylosing Spondylitis Quality of Life (ASQoL)

ASQoL up to week 20

For the change in ASQoL scores, between-treatment differences in the change in ASQoL scores will be evaluated using MMRM. Treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status and analysis visit will be used as categorical factors and baseline ASQoL score and weight as continuous covariates. Treatment by analysis visit and baseline ASQoL score by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the 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 and/or secukinumab at the appropriate analysis visits.

ASQoL at week 52

The change from baseline to Week 52 in ASQoL score will be analyzed by composite estimand strategy in which data after switch is set to extreme unfavorable value. Wilcoxon rank-sum test will be used for testing the difference in distributions of the composite endpoint in active vs placebo. Treatment differences will be summarised using trimmed means and confidence intervals will be calculated using permutation distribution.

[redacted]

SF-36

The following variables will be evaluated:

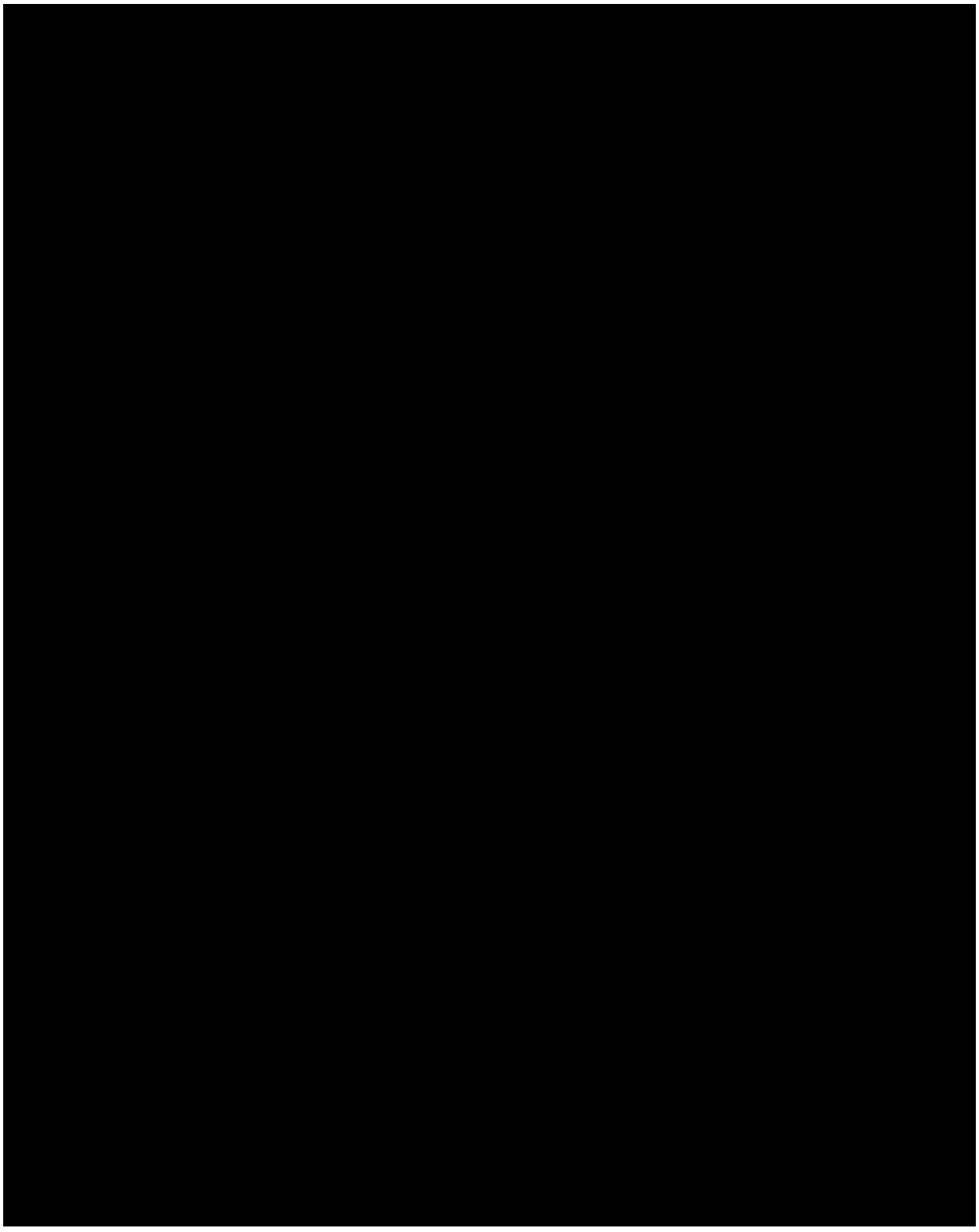
- SF-36 domain scores (based on a scale of 0-100).
- SF-36 PCS and MCS scores (norm-based scores).
- SF-36 PCS responder (improvement of ≥ 2.5 points, Lubeck 2004)

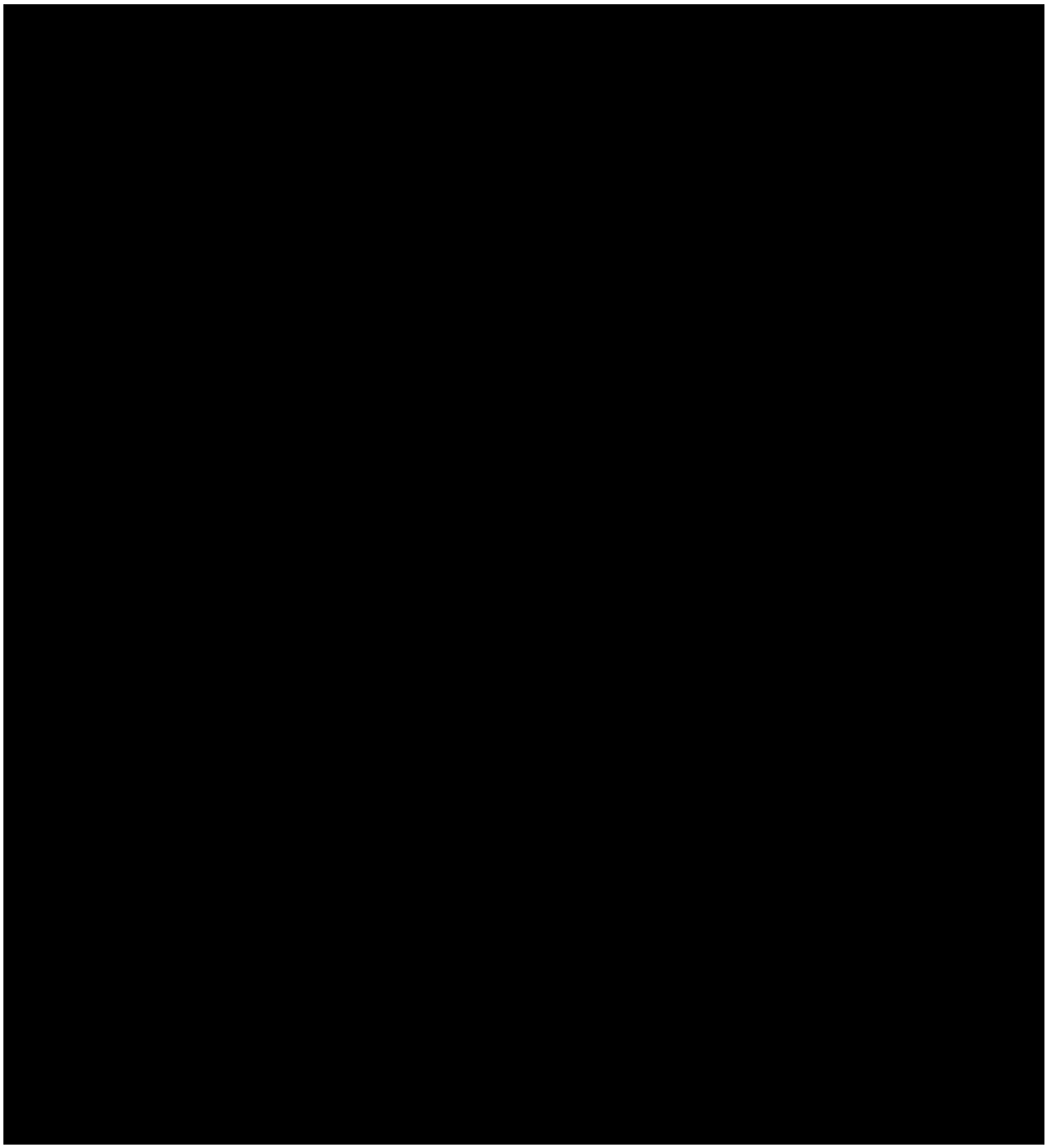
For the change in SF-36 summary scores (PCS and MCS), between-treatment differences will be evaluated using MMRM. Treatment group, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status and analysis visit will be included as categorical factors and baseline SF-36 score (PCS or MCS) and weight as continuous covariates. Treatment by analysis visit and baseline SF-36 score (PCS or MCS) by analysis visit will be included as interaction terms in the model. An unstructured covariance structure will be assumed for the model. Pairwise comparisons will be performed between secukinumab regimens and placebo and/or secukinumab at appropriate analysis visits.

In the responder analyses, treatment groups will be compared with respect to response to treatment using a logistic regression model with treatment, stratification factor (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+), TNF-α status as factors and baseline SF-36 score (PCS or MCS)

and weight as covariates. Odds ratios and 95% CI will be presented for appropriate treatment comparisons.

The SF-36 domain scores will be summarized by treatment.





14 Safety evaluation

Summaries may be performed separately for initial (Week 1-20) period and entire treatment period (including follow-up). Week 20 is chosen due to the fact that patients may be switched to another treatment as this time. Use of data up to and including the last visit before first switch opportunity provides an unbiased comparison between AIN and placebo; data collected beyond Week 20 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.

14.1 Adverse events

The crude incidence of treatment emergent adverse events (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 date + 84 days) will be summarized by primary system organ class and preferred term. Confidence intervals for the crude rate will be derived. In addition, exposure time-adjusted incidence rates including 95% confidence intervals will be provided for the entire treatment period. A graphical display 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, using a narrow search. The MedDRA version used for reporting the study will be described in a footnote.

Non-treatment emergent adverse events will be listed.

AEs after wash-out period(84 days after last dose of study drug) for standard of care treatment will be provided in a separate listing. Events will be reported in standard AE tables under randomized treatment according to rule of 84 days after last dose.

For 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 summary of those types of events as reported by the investigator and confirmed by adjudication will be provided.

An overview of the safety analyses which will be performed for treatment emergent AEs, on treatment labs, vital signs and ECG for each analysis period is described below.

Table 14-1 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 20	• crude incidence	• crude incidence	• crude incidence	• crude incidence	• crude incidence	• crude incidence
Entire Treatment	• crude incidence • exposure time adjusted incidence	• crude incidence	• crude incidence	• exposure time adjusted incidence	• crude incidence • exposure time adjusted incidence	• crude 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 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
- at Level 1 for Risks and SMQ analyses

If adjudication is performed, the adjudication events (myocardial infarction, stroke, and cardiovascular death) will be listed.

Algorithms for date imputations will be provided in Programming Datasets Specifications.

14.2 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.

Reported laboratory assessments with either a less than or greater than sign ("<" or ">") will be used for analysis after removal of the sign and conversion to standard unit. These laboratory data will be displayed in listings using the standard unit with the reported sign ("<" or ">").

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:

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: 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.

Table 14-2 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	Life-threatening consequences; urgent intervention indicated
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; >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 below:

Table 14-3 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 laboratory criteria**) 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:
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.

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

14.3 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 below:

Table 14-4 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

14.4 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). Both Bazett (QTcB) and Fridericia (QTcF) corrections will be presented for QTc.

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

- QTc > 500 msec
- QTc > 480 msec
- QTc > 450 msec
- QTc changes from baseline > 30 msec
- QTc 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.

14.6 Compound specific safety evaluation

Safety topics of interest, such as risks defined in the Safety Profiling Plan (SPP), Risk Management Plan or topics of interest regarding signal detection or routine analysis are defined in the Program Case Retrieval Sheet that is stored in CREDI at the path Cabinets/CREDI Projects/A/AIN457A/Integrated Medical Safety.

The crude incidence and exposure-adjusted incidence rates for SPP risks will be summarized. In addition, listings will be provided presenting which subjects experienced which risk.

Important note: For the evaluation of SPP risks primary and secondary system organ classes of the MedDRA dictionary will be considered.

15 Interim analyses

15.1 Analysis plan A

No formal interim analyses will be performed. The primary efficacy analyses will be performed after all patients complete the treatment period 1 of the study (Week 24 for analysis plan A to be used for e.g., EU file).

15.2 Analysis plan B

For analysis plan B week 52 endpoints will be tested at time of week 24 database lock.

It is estimated that at the time of Week 24 database lock approximately 70% of the subjects will have reached their Week 52 visit. The data needed to calculate the primary endpoint and the secondary endpoints (as outlined Analysis Plan B) will have been generated with the exception of SI joint edema on MRI (Week 52 MRI reading campaign will not be completed until all patients have completed Week 52). The primary endpoint is analyzed in the TNF-naive population which comprises at least 80% of the total population. Considering this, at least 105 subjects per treatment arm will have their primary endpoint evaluable at the time of Week 24 database lock. The expected sample size for the proposed Week 24 interim analysis of Analysis Plan B is provided in Table 15-1 Expected sample sizes per treatment arm and cumulative power for primary endpoint ASAS40. It is estimated that at the time of Week 24 interim analysis, 131 patients per treatment arm will have data available for the majority of Week 52 endpoints. Of the 131 patients, at least 105 are expected to be TNF-naive.

Table 15-1 Expected sample sizes per treatment arm and cumulative power for primary endpoint ASAS40

	Population			
	TNF naive		All	
Analysis	Sample size	Cumulative power for ASAS40	Sample size	Cumulative power for ASAS40
Interim analysis	105	78%	131	91%
Final analysis	148	91%	185	99%

A Lan-De Mets alpha spending function with O'Brien Fleming type stopping boundary (as implemented in the software East 6.3) will be used to maintain the overall type-I error rate for the primary and secondary endpoints at week 52.

Based on the choice of α-spending function described above, the efficacy boundary in terms of p-value scale at the interim analysis is calculated as p=0.016 for a two-sided test. The observed (i.e., nominal) p-value has to be smaller than 0.016 to conclude superior efficacy at the interim analysis. If a hypothesis is not rejected at the interim analysis, it will be tested again at the Week 52 analysis after the Week 52 database lock, when the full Week 52 data are available. The efficacy boundary for the final analysis in terms of p-value scale is p=0.046.

The exact rejecting boundaries will be calculated after the exact number of patients in each treatment arm is available.

A sequential testing hierarchy will be used to test the secondary hypotheses. The secondary hypotheses will be tested at the interim analysis only if the primary hypothesis is rejected. This guarantees the 5% overall level significance for the primary and secondary hypotheses (Glimm 2010). For Week 52 endpoints, the same rejecting boundaries will be used as for the primary endpoint. For Week 16 endpoints, the full data will be available at the time of the interim analysis, and they will be tested with full alpha at the Week 24 interim analysis.

Table 15-2 shows all hypothesis tested at interim analysis and the expected rejection boundaries for each hypothesis.

Table 15-2 Group sequential testing boundaries for each hypothesis

Hypothesis	Dose	Endpoint	Proportion of data available at interim analysis	Efficacy boundaries on p-value scale[1]	
				Interim analysis	Final analysis
H_1	secukinumab 150 mg (without load)	ASAS40 in TNF naïve at Week 52	71%	0.016	0.046
H_2	secukinumab 150 mg (without load)	ASAS40 at Week 52	71%	0.016	0.046
H_3	secukinumab 150 mg (without load)	ASAS40 at Week 16	100%	0.05[2]	0.05[2]
H_4	secukinumab 150 mg (with load)	ASAS40 in TNF naïve at Week 52	71%	0.016	0.046
H_5	secukinumab 150 mg (with load)	ASAS40 at Week 52	71%	0.016	0.046
H_6	secukinumab 150 mg (with load)	ASAS40 at Week 16	100%	0.05[2]	0.05[2]
H_7	secukinumab 150 mg (without load)	change from baseline in total BASDAI at Week 16	100%	0.05[2]	0.05[2]
H_8	secukinumab 150 mg (without load)	BASDAI50 at Week 16	100%	0.05[2]	0.05[2]
H_9	secukinumab 150 mg (without load)	BASDAI50 at Week 52	71%	0.016	0.046
H_{10}	secukinumab 150 mg (without load)	change from baseline in hsCRP at Week 16	100%	0.05[2]	0.05[2]

H_{11}	secukinumab 150 mg (without load)	change from baseline in SF-36 PCS at Week 16	100%	0.05[2]	0.05[2]
H_{12}	secukinumab 150 mg (without load)	change from baseline in ASQoL at Week 16	100%	0.05[2]	0.05[2]
H_{13}	secukinumab 150 mg (without load)	ASAS5/6 at Week 16	100%	0.05[2]	0.05[2]
H_{14}	secukinumab 150 mg (without load)	ASAS20 at Week 16	100%	0.05[2]	0.05[2]
H_{15}	secukinumab 150 mg (without load)	change from baseline in BASFI at Week 16	100%	0.05[2]	0.05[2]
H_{16}	secukinumab 150 mg (without load)	change from screening in SI joint edema on MRI at Week 16	100%	0.05[2]	0.05[2]
H_{17}	secukinumab 150 mg (without load)	ASDAS-CRP ID at Week 52	71%	0.016	0.046
H_{18}	secukinumab 150 mg (without load)	change from baseline in ASQoL at Week 52	71%	0.016	0.046
H_{19}	secukinumab 150 mg (without load)	change from screening in SI joint edema on MRI at Week 52	0%	NA	0.05
H_{20} to H_{32}	These hypotheses appear after after H_{19} in the testing hierarchy and will not be tested at interim analysis		NA	NA	0.05

[1]These are estimated rejection boundaries based on expected sample sizes on each analysis. The rejection boundaries will be updated after exact number of observation is available.

[2]Week 16 endpoints have full data available at interim analysis and therefore they are tested with full alpha at both analysis. They will be tested at interim analysis if all week 52 hypotheses higher in the hierarchy have been rejected, otherwise they will be tested at the final analysis.

15.3 Confidentiality of week 24 interim analysis results

Although unblinding will occur after the Week 24 database lock, the original randomization to active treatment vs placebo will continue to remain blinded to all investigators, site personnel and patients until all patients have completed the treatment period 2 (Week 52) and the Week 52 database lock has occurred. A separate study team will conduct the 24 Week interim analysis and no access to the interim results or individual treatment assignments will be given to the study team conducting the ongoing trial until the Week 52 database lock.

To ensure consistent efficacy result from the first 24 weeks, efficacy analyses up to Week 20 related to the primary and secondary endpoints will not be re-analyzed for the next lock at Week 52 as part of the statistical hierarchy.

Example of how this will be handled in practice:

- For ASAS40 analysis at Week 16 (which is part of both statistical testing hierarchies) the result will be made consistent across analysis plans by using the value from the Week 24 lock even if the data slightly changes for the Week 52 lock. Re-analyzed values related to

the primary endpoint may be presented in the statistical appendix for comparative purposes.

16 Determination of sample size

In a study on the same indication of similar design (Landewé et al 2014) an ASAS40 response rate of 47.1% for active treatment and 16% for placebo was observed at Week 12, however this trial had a limited number of TNFα-IR patients and a meta-analysis (MA) from unpublished studies with secukinumab in AS indicates that the placebo rates observed in recent AS studies may be higher. Hence, assumptions are based on the result of active treatment from this TNF-inhibitor study in nr-axSpA (but adjusted for the expected inclusion of TNF-IR patients) and with placebo response rates taken from the secukinumab MA. This MA included approximately 25% TNF-IR patients, and the ASAS40 response rate in the 150 mg dose for TNF-IR was 76% of the response in the TNF naïve group. Assuming 20% of randomized patients will be TNF-IR and have the same TNF-IR vs TNF naïve response ratio as seen in the MA (76%), the estimate for the entire population is 44.8% (i.e., 47.1%*0.8 + 47.1%*0.2*0.76) for secukinumab and 25.9% for placebo. ASAS40 in TNF naïve patients only is assumed to be 47.1% for secukinumab and 27.9% for placebo.

At Week 16, response rates for the secukinumab 150 mg s.c. regimens with loading and without loading are assumed to be the same.

Primary endpoint for analysis plan A

An overall type I error (2-sided) 5% will be used to control type I error. Since the hierarchy is sequential starting with secukinumab with load tested versus placebo the full type I error will be utilized for each comparison.

Table 16-1 Summary of power for binary primary endpoint

Endpoint	Placebo Response Rate	Secukinumab with SC Loading		Secukinumab without SC Loading	
		Response Rate	Power N=185/arm	Response Rate	Power N=185/arm
ASAS40 TNF naïve at Week 16	27.9%	47.1%	91%	47.1%	91%

Based on these assumptions including 185 patients per arm would give 91% power to reject a hypothesis of equal response rate based on Fisher's exact test. Note that the power of 91% is for the subgroup of TNF naïve (185*0.8=148), but to achieve 90% power for important secondary endpoints 185 patients will be included per arm.

Primary endpoint for analysis plan B

An overall type I error (2-sided) 5% will be used to control type I error. Since the hierarchy is sequential starting with secukinumab without load tested versus placebo the full type I error will be utilized for each comparison.

Based on an assumption that the ASAS40 TNF naïve response rate at Week 16 will be 47.1% for secukinumab and 27.9% for placebo, that 80% of secukinumab patients and 65% of placebo patients will retain their response status from Week 16, and that 10% of secukinumab patients and 5% of placebo patients will be new responders at Week 52, the ASAS40 response rate at Week 52 based on non-responder imputation can be estimated as 43.0% and 21.7%, respectively, resulting in a power of 97% to reject an hypothesis of equal response rate at the 5% level.

Table 16-2 Summary of power for binary primary endpoint

Endpoint	Placebo Response Rate	Secukinumab with SC Loading		Secukinumab without SC Loading	
		Response Rate	Power N=185/arm	Response Rate	Power N=185/arm
ASAS40 TNF naïve at Week 52	21.7%	43.0%	97%	43.0%	97%

Secondary endpoints for analysis plan A

In Table 15-3 the power for rejecting each of the binary secondary hypotheses at Week 16 is presented based on Fisher's exact test with a 2-sided 5% type I error rate and 185 patients per group assuming the response rates without SC loading are the same as for with load.

Table 16-3 Power for binary secondary variables at Week 16

Endpoint	Placebo Response Rate	Secukinumab with SC Loading		Secukinumab without SC Loading	
		Response Rate	Power N=185/arm	Response Rate	Power N=185/arm
ASAS40	25.9%	44.8%	96%	44.8%	96%
ASAS5/6	24.3%	45.0%	98%	45.0%	98%
BASDAI50	19.4%	45.4%	99%	45.4%	99%
ASAS20	43.7%	61.1%	90%	61.1%	90%
ASAS PR[1]	8.7%	21.9%	93%	21.9%	93%

Source: Active from R Landewé et al 2014 (Week 12 result, 400 mg Q4W) adjusted for TNF-IR and Placebo from a meta-analysis of unpublished Novartis AS studies (CAIN457F2314 and CAIN457F2320). Calculations based on Fisher's exact test (nQuery 7.0 PTT2-1)

[1]Week 16 response for TNF naïve patients assumed 25% for secukinumab and 10% for placebo which were then adjusted for 20% TNF-IR.

In Table 15-4 the power for rejecting each of the secondary hypotheses at Week 16 based on continuous variables are presented with assumed change from baseline and standard deviation (SD). Following the assumption for response variables the change from baseline was adjusted for inclusion of TNFα-IR patients but with unchanged SD. The change from baseline without SC loading is assumed the same as for with load.

Table 16-4 Power for continuous secondary variables at Week 16

Endpoint	Common SD	Placebo Mean	Secukinumab with SC Loading	Secukinumab without SC Loading

			Mean	Power N=185/arm	Mean	Power N=185/arm
BASDAI[1]	2.3	-1.75	-3.30	99%	-3.30	99%
hsCRP[2]	0.867	0.095	-0.544	99%	-0.544	99%
BASFI[1]	2.1	-1.37	-2.20	96%	-2.20	96%
MRI SI joint edema[3]	2.17	-0.17	-1.30	99%	-1.30	99%
SF-36 PCS[4]	7.15	3.67	6.32	94%	6.32	94%
ASQoL[4]	4.55	-2.32	-4.10	96%	-4.10	96%

SD=standard deviation. Result is mean change from baseline with CRP in log scale.

[1] Active from R Landewé et al 2014 (Week 12 result, 400 mg Q4W) adjusted for TNF-IR and Placebo from a meta-analysis of unpublished AS studies (CAIN457F2314 and CAIN457F2320).

[2] Meta-analysis of unpublished Novartis AS studies (CAIN457F2314 and CAIN457F2320)

[3] Novartis AS study (AIN457F2305)

[4] Meta-analysis of Novartis AS studies (CAIN457F2310 and CAIN457F2320)

Calculations based on two-sample t-test (nQuery 7.0 MTT0-1)

Secondary endpoints for analysis plan B

Response rates for load and without load at weeks 16 and 52 are assumed to be the same.

Table 16-5 Power for binary secondary variables

Endpoint	Placebo Response Rate	Secukinumab with SC Loading		Secukinumab without SC Loading	
		Response Rate	Power N=185/arm	Response Rate	Power N=185/arm
ASAS40 wk52	20.5%	41.4%	99%	41.4%	99%
ASAS40 wk16	25.9%	44.8%	96%	44.8%	96%
BASDAI50 wk16	19.4%	45.4%	99%	45.4%	99%
BASDAI50 wk52	16.6%	41.8%	99%	41.8%	99%
ASAS5/6 wk16	24.3%	45.0%	98%	45.0%	98%
ASAS20 wk16	43.7%	61.1%	90%	61.1%	90%
ASDAS ID wk52[1]	10.2%	25.3%	96%	25.3%	96%

Week 16

Source: Active from R Landewé et al 2014 (Week 12 result, 400 mg Q4W) adjusted for TNF-IR and Placebo from a meta-analysis of unpublished Novartis AS studies (CAIN457F2314 and CAIN457F2320).

Week 52

Estimates from Week 16 were adjusted for assumed response retention.

[1] Week 16 response for TNF naïve patients assumed 25% for secukinumab and 10% for placebo which were then adjusted for 20% TNF-IR and assumed response retention.

ASDAS ID = ASDAS CRP Inactive Disease (ASDAS-CRP score < 1.3)

Calculations based on Fisher's exact test (nQuery 7.0 PTT2-1)

Table 16-6 Power for continuous secondary variables

Endpoint	Common SD	Placebo Mean	Secukinumab with SC Loading		Secukinumab without SC Loading	
			Mean	Power N=185/arm	Mean	Power N=185/arm
BASDAI wk16[1]	2.30	-1.75	-3.30	99%	-3.30	99%
hsCRP wk16[2]	0.867	0.095	-0.544	99%	-0.544	99%
SF-36 PCS wk16[3]	7.15	3.67	6.32	94%	6.32	94%

Endpoint	Common SD	Placebo Mean	Secukinumab with SC Loading		Secukinumab without SC Loading	
			Mean	Power N=185/arm	Mean	Power N=185/arm
BASFI wk16[1]	2.10	-1.37	-2.20	96%	-2.20	96%
MRI SI joint edema wk16[4]	2.17	-0.17	-1.30	99%	-1.30	99%
ASQoL wk16[3]	4.55	-2.32	-4.10	96%	-4.10	96%

SD=standard deviation. Result is mean change from baseline with CRP in log scale.

Week 16

[1]Active from R Landewé et al 2014 (Week 12 result, 400 mg Q4W) adjusted for TNF-IR and Placebo from a meta-analysis of unpublished Novartis AS studies (CAIN457F2314 and CAIN457F2320).

[2]Meta-analysis of unpublished Novartis AS studies (CAIN457F2314 and CAIN457F2320)

[3]Meta-analysis of Novartis AS studies (CAIN457F2310 and CAIN457F2320)

[4]Novartis AS study (CAIN457F2305).

Power for continuous composite estimands depends on the difference in continuous endpoint and difference in treatment escape rates. Assuming the change from baseline being the same at Week 52 as at Week 16 and no escape, the power for Wilcoxon rank sum test would be similar to power of two-sample t-test (ASQoL 96% and MRI SI joint edema 99%). If escape rates are similar in both treatment arms the power will decrease if overall escape rate is increasing, whereas the power will remain the same or increase if escape rates are different between treatment arms.

Appendix

17.1 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, 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 below.

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 17-1 Analysis visit windows

Analysis Visit		Target Day	Analysis Visit Window	Group 1	Group 2	Group 3	Group 4	[redacted]	Group 6	Group 7	Group 8	[redacted]
Baseline		1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1		≤ 1	≤ 1	≤ 7	
Week 1		8	2-11	2-11	2-11	2-11						
Week 2		15	12-18	12-18	12-18	12-22						
Week 3		22	19-25	19-25	19-25							
Week 4		29	26-43	26-43	26-43	23-43						
Week 8		57	44-71	44-71	44-71	44-71	2-85		2-85			
Week 12		85	72-99	72-99	72-99	72-99						
Week 16		113	100-127	100-127	100-127	100-141	86-141		86-239	2-239	8-239	
Week 20		141	128-155	128-155	128-155							
Week 24		169	156-183	156-183	156-183	142-183	142-267					
Week 28		197	184-211	184-211	184-239	184-239						
Week 32		225	212-239	212-239								
Week 36		253	240-267	240-267								
Week 40		281	268-295	268-295	240-323	240-323						
Week 44		309	296-323	296-323								
Week 48		337	324-351	324-351								
Week 52		365	352-379	352-407	324-407	324-407	268-449		240-547	240-547	240-547	
Week 56		393	380-407									
Week 60		421	408-435									
Week 64		449	436-463	408-491	408-491	408-491						
Week 68		477	464-491									
Week 72		505	492-519									

Analysis Visit		Target Day	Analysis Visit Window	Group 1	Group 2	Group 3	Group 4	[redacted]	Group 6	Group 7	Group 8	[redacted]
Week 76		533	520-547	492-575	492-575	492-575	450-631					
Week 80		561	548-575									
Week 84		589	576-603									
Week 88		617	604-631	576-673	576-673	576-673						
Week 92		645	632-659									
Week 96		673	660-687									
Week 100		701	688-715									
Week 104		729	716-743	674-743	674-743	674-743	632-743		548-743	548-743	548-743	
	Group 1: Patient's global assessment of disease activity (VAS), Patient's assessment of back pain intensity (VAS), BASFI, BASDAI, [redacted], [redacted] hsCRP Group 2: Vital signs Group 3: Hematology, blood chemistry, urinalysis Group 4: [redacted], ASQoL, SF-36, [redacted] [redacted] Group 6: Lipids Group 7: Weight, ECG Group 8: MRI [redacted]											

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.

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 17-2 Rules for flagging variables

Timing of measurement	Type of data	Rule
Baseline	All data	Baseline is defined as the last available measurement recorded on or before the reference start date (only date part)/Day 1. If there are multiple assessments on Day 1, the following rules apply: 1. 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 2. If assessment time does not exist: select the available measurement from the lowest CRF visit number
Post-baseline efficacy	All data	The measurement closest to the target day will be used. In the event two measurements are taken equally apart (e.g. 1 day before target date and 1 day after) the first one will be used. If the patient switches treatment (e.g. from placebo to AIN) within the window the following rules apply: • If available, the closest measurement to the target date which is on or before the switch date will be used • If there are no data on or before the switch date then the closest measurement after the switch to target will be used Cases where the same parameter is recorded more than once on the same date will be handled as follows: • If time of completion exists the earliest measurement will be used • 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, ECG, etc.)	The measurement closest to the target day will be used. In the event two measurements are taken equally apart (e.g. 1 day before target date and 1 day after) the first one will be used. If the patient switches treatment (e.g. from placebo to AIN) within the window the following rules apply: • If available, the closest measurement to the target date which is on or before the switch date will be used • If there are no data on or before the switch date then the closest measurement after the switch to target will be used Cases where the same parameter is recorded more than once on the same date will be handled as follows: • If time of completion exists the earliest measurement will be used • 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

17.2 Statistical methodology and assumptions

17.2.1 Analysis of continuous data

17.2.1.1 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.

17.2.1.2 Analysis of covariance

Univariate model

An analysis of covariance (ANCOVA) model will be used to analyze some endpoints. The model will include factors and covariates as specified for respective analysis. The SAS code below outlines a template for the analysis where covariates can be added or removed as required:

ods output diffs=*lsm_diff* lsmeans=*lsm_est*;

proc mixed data=;

by *visit*;

class *treatment strata tnf_status*;

model *outcome = treatment strata tnf_status baseline weight*;

lsmeans *treatment* / diff cl;

run;

Least-square-mean (LSM) estimates for each treatment group and LSM difference, confidence intervals and p-value for the difference between each dose of secukinumab and placebo, and between secukinumab doses if relevant, can be obtained.

Repeated measures analysis

Some endpoints will be analyzed using a longitudinal model that comprises several visits. The model used will be mixed model repeated measures (MMRM) with factors, covariates, interactions and covariance structure as specified for respective analysis. The SAS code below outlines a template for the analysis where covariates and interaction terms can be added or removed as required:

ods output diffs=*lsm_diff* lsmeans=*lsm_est*;

proc mixed data=;

class *treatment strata visit tnf_status*;

model *outcome = treatment strata visit tnf_status baseline weight treatment*visit baseline*visit* / ddfm=kr;

repeated *visit* / type=un subject=;

lsmeans *treatment*visit* / diff cl;

run;

Least-square-mean (LSM) estimates for each treatment group and LSM difference, confidence intervals and p-value for the difference between each dose of secukinumab and placebo, and between secukinumab doses if relevant, will be calculated at appropriate analysis visits.

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

1. change to ddfm=bw. If still no convergence, perform step 2.

2. change to type=cs. If still no convergence, perform step 3.

3. remove covariates in the following order until convergence: *weight*, *baseline*visit*, *TNF-α inhibitor status*, *strata*.

17.2.2 Analysis of binary and categorical data

17.2.2.1 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-α/2)-quantile of the standard normal distribution (SAS: z=probit(1-α/2), n as total number of subjects (i.e., number of subjects in the denominator), p as estimated crude incidence (number of subjects with event / n) and 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

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

Note: if $p < L$ then $L = p$ and if $p > U$ then $U = p$.

17.2.2.2 Logistic regression

Certain binary outcome variables, e.g., response outcomes, will be evaluated using a logistic regression model. The model will include factors and covariates as specified for respective analysis. The SAS code below outlines a template for the analysis where covariates can be added or removed as required:

ods output diffs=*lsm_diff* convergencestatus=*conv_status*;

proc logistic data=;

by *visit*;

class *treatment strata tnf_status* / param=glm;

model *response = treatment strata tnf_status weight*;

lsmeans *treatment* / diff cl exp;

run;

For cases where the convergence status indicates that the model did not reach appropriate convergence (*conv_status* is not 0), no odds ratio or p-value will be presented from that model. However, if the issue relates to the primary timepoint of Week 16 or Week 52 then the following steps will be followed:

1. Remove TNF status from the model. If there are still issues perform step 2

2. Run an exact logistic model as described below

In cases where separation is a concern for the primary timepoint Week 16 or Week 52, e.g., 0% or 100% response for some treatment and covariate level combination, an exact logistic regression model may be applied to all visits. To help with convergence, this model will not include any continuous covariates. The SAS code below outlines a template for the analysis where covariates can be added or removed as required:

ods output exactoddsratio=*exact_or*;

proc logistic data=*data* exactonly;

by *visit*;

class *treatment strata tnf_status* / param=ref;

model *response = treatment strata tnf_status*;

exact *treatment* / estimate=odds;

run;

When exact logistic regression is unable to be implemented (due to computational complexity as the procedure can lead to extremely long run times), then Fisher's exact test will be applied. In this case, only a p-value for a test of equal response in the two groups can be obtained (no odds ratios or confidence intervals can be estimated.)

ods output fishersexact=*fisher*;

proc freq data=;

by *visit*;

table *treatment***response* / fisher;

run;

Input dataset should only contain data from the two treatment groups to be compared.

17.2.3 Imputation methods

17.2.3.2 Multiple Imputation

Under MAR assumption

A linear regression model will be used to perform multiple imputation (MI) under a missing-at-random (MAR) assumption. To help preserve the relationship between outcome and covariates within each treatment a separate model will be run for each treatment. This will also help ensure that the imputation model does not make stronger assumptions on data relations than the analysis model.

The SAS code below outlines a template for the analysis where covariates and visits can be added or removed as required. To ensure that results can be replicated the data should be sorted by subject number before running the model (the data should be in horizontal format with one subject per dataset row).

proc mi data= seed=4572315 nimpute=100 out=*mi_out*;

by *treatment*;

class *strata tnf_status*;

fcs reg (/details);

var *strata tnf_status weight value1 value2 value3*;

run;

Where in the template code the continuous variable to be imputed is *value* (e.g., *value1* could be the baseline value and *value2* the first post-treatment measurement of the variable to be imputed.) Normally, all data collection visits during the analysis period of interest would be included in the model. Including variables using a CLASS statement instead of a BY statement

should help facilitate model convergence also when the number of non-missing data points are low for some specific covariate level and visit combination. The FCS option is used to ensure that also non-monotone missing data can be handled in an appropriate way.

If convergence is not obtained the following sequential steps will be used:

1. remove *weight*. If still no convergence, perform step 2

2. remove *TNF-α inhibitor status*.

For a situation where several variables need to be imputed using separate models (e.g., using independent models to impute each component needed to derive a response variable *V*) a step-wise process needs to be implemented as outlined below:

1. Run the SAS code as described above for the first variable to be imputed
2. Run the SAS code as described above for the next variable to be imputed (but with the following changes: "data=*mi_out*", "out=*mi_out2*", "by *treatment _imputation_*", "nimpute=1")
3. Repeat step 2, but with input dataset equal to the output dataset from the prior step, until all *j* variables have been imputed resulting in a dataset named mi_out_j
4. Derive the variable *V* from within mi_out_j

The required analysis (e.g., ANCOVA) is then performed separately within each imputation dataset (as identified by variable *_imputation_*). To obtain the final result of the imputation process the analysis result from each imputation dataset needs to be combined according to Rubin's rules as outlined below:

ods output parameterestimates=*mi_result*;

proc mianalyze data=;

modeleffects *estimate*;

stderr *estimate_se*;

run;

The *estimate* and *estimate_se* parameters come from the analysis model used to analyze the imputed variable within each imputation dataset (e.g., from the lsmean estimate of the treatment difference and its standard error obtained from PROC LOGISTIC or PROC MIXED.)

To obtain binary response rates and confidence intervals for individual treatment groups the following process should be followed (exemplified for one visit):

ods output binomialprop=*bin_est*;

proc freq data=;

by *treatment _imputation_*;

table *response* / binomial (cl=wilson correct);

run;

Then apply a logit transformation on the saved proportions and derive its standard error:

data *bin_est*; set *bin_est*;

estimate=log(*_bin_*/(1-*_bin_*));

estimate_se=*e_bin*/(*_bin_**(1-*_bin_*));

run;

The transformed binomial proportion estimates and its standard errors are then combined by applying Rubin's rules as described above using PROC MIANALYZE. Before presenting the combined data it needs to be transformed back as follows:

data *mi_result*; set *mi_result*;

prop_est=1/(1+exp(-*estimate*));

prop_lower=1/(1+exp(-*lclmean*));

prop_upper=1/(1+exp(-*uclmean*));

run;

If all responses are imputed the same as 0 (or 1) for all imputation datasets for a specific treatment or subgroup then the between-imputation-variation will be zero. The combined final response rate for the specific treatment or subgroup would be presented as seen in any of the imputed datasets together with 95% CI from Wilson's method (as obtained from PROC FREQ).

The following steps will be performed to handle special cases:

- If after imputation all responses (observed+imputed) are the same either 0 or 1 for all imputation datasets for a specific treatment or subgroup it will not be possible to perform a logit transformation and the response rate (0% or 100%) for these cases will be presented together with the 95% CI from Wilson's method (as obtained from PROC FREQ).
- If after imputation the average response rate is the same across all imputed datasets (but not 0 or 1) there is no between dataset variation and Rubin's rules cannot be applied. Instead the average response will be used with 95% CI from Wilson's method (as obtained from PROC FREQ)

17.2.4 Crude incidence and related risk estimates

17.2.4.2 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 can be obtained by using the SAS procedure PROC FREQ with statement EXACT OR. However, to be able to adjust for covariates odds ratios will primarily be obtained from PROC LOGISTIC.

[redacted]

17.2.5 Exposure adjusted incidence rate and related risk estimates

17.2.5.1 Exposure adjusted incidence rate and 100*(1-α)% confidence interval

It will be assumed that for each of n subjects in a clinical trial the time t_j (j=1,...,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 θ. The rate parameter θ 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-α)% confidence interval for a Poisson variable with parameter θT and observed value D can be obtained based on (Garwood 1936), from which an exact 100*(1-α)% confidence interval for D/T will be derived as follows (Sahai 1993; Ulm 1990):

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

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

Where $c_{\alpha,k}$ is the αth 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 17-4 Examples for calculating exposure time for incidence rates (IR)

1st treatment / total exposure time	2nd treatment / total exposure time	AE event onset (in days from study start)	Exposure for IR
Placebo / 100 days	AIN457 150 mg / 200 days	Day 50 (during 1st treatment) Day 110 (10 days into 2nd treatment)	Placebo: 50 days AIN457 150 mg: 10 days Any AIN: 10 days

17.2.7 RANK based analysis

Rank based test using Wilcoxon rank sum test will be used for composite continuous estimands. A SAS code template for hypothesis testing for treatment1 vs treatment2 is given below:

proc npar1way data= wilcoxon;

where *treatment* in ("*treatment1*", "*treatment2*");

class *treatment*;

var *rankvariable*;

run;

The analysis variable *rankvariable* will either be the based on the observed value or extreme unfavorable value, if patient has discontinued, has missing data or has switched treatment.

17.2.8 Trimmed means

Trimmed means will be calculated using methodology presented in (Permutt 2017). Trimming is compatible with a wide range of analytical methods. The trimming will be combined with the appropriate analytical method(e.g. MMRM or ANCOVA). The general algorithm can be stated simply:

1. Order the data and trim equal fractions, always trimming all dropouts.
2. Compute a summary measure on the trimmed data.
3. Repeat steps 1 and 2 on permuted data to construct a reference distribution for testing and interval estimation

The fraction trimmed will be chosen adaptively as the greatest of the proportions of dropouts in the treatment groups to be compared.

[redacted]

17.3 Group sequential design used in phase III studies

The statistical methodology for the interim analyses of week 52 endpoints will be based on group sequential methodology with efficacy stopping boundaries defined by type I error spending functions.

This approach is flexible in dealing with any deviations from the targeted sample sizes, or unexpected changes to the plan.

If the exact number of subjects observed at the interim and final analyses deviates from the target numbers described in the statistical analysis plan, the actual critical boundaries will be derived using the prespecified

error spending functions and the actual numbers of subjects observed.

- At interim analysis, information fractions will be computed as the ratio of the number of observations at the interim analysis relative to the number targeted for the final analysis, as described in the sample size section of the protocol.
- At the final analysis, the critical value will be calculated using the exact number of observations at the final analysis, considering the α-levels spent at interim analyses and considering the actual correlation among the test statistics, in order to achieve a cumulative type I error smaller than the desired significance level (i.e., smaller than 2.5% for a one-sided test and smaller than 5% for a two-sided test).

17.3.1 Alpha spending function

The stopping boundaries to be used for week 52 tests will be calculated using the α-spending function approach described in Lan and DeMets (Lan and DeMets, 1983). The spending function for one-sided test has the following functional form:

$$\alpha(t) = 2 - 2\Phi(\, Z_{\alpha/2} \,/\sqrt{t}\,)$$

This function generates stopping boundaries that closely resemble the O'Brien-Fleming boundaries (O'Brien and Fleming, 1979).

18 References

Braun J, McHugh N, Singh A, Wajdula JS, Sato R (2009) Improvement in patient-reported outcomes for patients with ankylosing spondylitis treated with etanercept 50 mg once-weekly and 25 mg twice-weekly. Rheumatology (Oxford); 46: 999-1004

Bretz F, Maurer W, Brannath W, et al. (2009) A graphical approach to sequentially rejective multiple test procedures. Stat Med; 28(4):586-604.

James R. Carpenter , James H. Roger & Michael G. Kenward (2013) Analysis of Longitudinal Trials with Protocol Deviation: A Framework for Relevant, Accessible Assumptions, and Inference via Multiple Imputation, Journal of Biopharmaceutical Statistics, 23:6, 1352-1371

[redacted]

Corticosteroid converter (data retrieved December 12, 2013):
http://www.globalrph.com/steroid.cgi

Davis JC, Revicki D, van der Heijde D, et al. (2007) Health-related quality of life outcomes in patients with active ankylosing spondylitis treated with adalimumab: results from a randomized controlled study. Arthritis & Rheumatism; 57(6):1050-1057.

Doward LC, Spoorenberg A, Cook SA, et al. (2003) Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. Ann Rheum Dis; 62(1):20-26.

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

Glimm, E., Maurer, W., and Bretz, F. (2010) Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine 29, 219-228

Haywood KL, et al. (2002) Disease-specific, patient-assessed measures of health outcome in ankylosing spondylitis: reliability, validity and responsiveness. Rheumatology; 41:1295-1302.

Huffman K, Samsa G, Slentz C, et al. (2006) Response of high-sensitivity C-reactive protein to exercise training in an at-risk population. Am Heart J; 152(4):793-800.

Inman RD, David JC Jr, et al. (2008) Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. Arthritis Rheum; 58(11):3402-12.

Lan, K. K. G. & Demets, D. L. (1983). Discrete sequential boundaries for clinical trials.

Biometrika 70, 659-63

Landewé R, Braun J, Deodhar A, et al. (2014) Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis 2014; 73:39–47

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

Lukas C, Landewé R, Sieper J. (2009) Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis.; 68(1):18-24.

Machado PM, Landewé RB, van der Heijde DM. (2011) Endorsement of definitions of disease activity states and improvement scores for the ankylosing spondylitis Disease Activity Score: results from OMERACT 10. J Rheum; 38(7):1502-1506.

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

O'Brien PC, Fleming TR (1979). A multiple testing procedure for clinical trials, Biometrics.

35(3):549-56.

Permutt T, Li F(2017). Trimmed means for symptom trials with dropouts. Pharmaceutical statistics.;16(1):20-8.

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

Sieper J, Rudwaleit M, Baraliakos X. (2009) The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis.; 68 Suppl 2:ii1-44.

Sieper J, et al. (2013) Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72:815–822.

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 D, Dijkmans B, et al. (2005) Efficacy and safety of infliximab in patients with ankylosing spondylitis: Results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum; 52(2):582-591.

van der Heijde D, Kivitz A, Schiff MH, et al. (2006) Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheum; 54(7):2136-46.

[redacted]

van Tubergen A, Landewe R, van der Heijde D, et al. (2001) Combined spa-exercise therapy is effective in patients with ankylosing spondylitis: a randomized controlled trial. Arthritis Rheum; 45: 430–38

Yellen SB, Cella DF, Webster K, et al. (1997) Measuring Fatigue and Other Anemia-Related Symptoms with the Functional Assessment of Cancer Therapy (FACT) Measurement System. Journal of Pain and Symptom Management; 13:63-74.

[redacted]

NOVARTIS

Clinical Development

Secukinumab (AIN457)

CAIN457H2315 / NCT02696031

A randomized, double-blind, placebo-controlled multicenter study of secukinumab 150 mg in patients with active non-radiographic axial spondyloarthritis to evaluate the safety, tolerability and efficacy up to 2 years, followed by an optional phase of either 150 mg or 300 mg randomized dose escalation for up to another 2 years

Statistical Analysis Plan (SAP) – Extension Phase Amendment 1

Author: [redacted], Trial Statistician

Document type: SAP Documentation

Document status: Final

Release date: 09-Apr-2021

Number of pages: 36

Document History – Changes compared to previous version of SAP.

Version	Date	Changes
1	19-Jan-2021	Initial version
Amendment 1	08-April-2021	- Add clarification about analysis set, since there are subjects who were randomized before ASAS20 response status at Week 104 was confirmed. - Update the wording about COVID-19 related PDs. - Remove the subgroup analysis on TNF-α inhibitor status and stratification factor, since it is not planned for the final CSR. - Remove the paragraphs about prior nr-axSpA therapy, since it is not of interest for the extension phase. - Remove the paragraphs about lipids, since it is not of interest for the extension phase. - Add clarification about analysis visit. - Update the sample code of MI to solve the convergence issue.

Table of contents

List of tables

List of abbreviations

AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AS	Ankylosing Spondylitis
ASAS	Assessment of Spondyloarthritis International Society criteria
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASQoL	Ankylosing Spondylitis Quality of Life Questionnaire
AST	Aspartate Aminotransferase
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BMI	Body Mass Index
BSL	Baseline
CRF	Case Report/Record Form (paper or electronic)
CRP (hsCRP)	(high sensitivity) C-Reactive Protein
CRP+	hsCRP > ULN (as defined by the central lab)
CTCAE	Common Criteria for Adverse Events
DMARD	Disease Modifying Antirheumatic Drug
DMC	Data Monitoring Committee
EAS	Extension Phase Full Analysis Set
ECG	Electrocardiogram
ESR	Erythrocyte Sedimentation Rate
FACIT-Fatigue	Functional Assessment of Chronic Illness Therapy – Fatigue
FDA	Food and Drug Administration
GGT	Gamma Glutamyl Transferase
HDL	High Density Lipoprotein
HLA	Human Leukocyte Antigen
IRT	Interactive Response Technology
ITT	Intent-to-Treat
LDL	Low Density Lipoprotein
LFT	Liver Function Test
LLN	Lower Limit of Normal

LOCF	Last Observation Carried Forward
LSM	Least-Square-Mean
MAR	Missing At Random
MASES	Maastricht Ankylosing Spondylitis Enthesis Score
MCS	Mental Component Summary
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram(s)
MI	Multiple Imputation
MMRM	Mixed effect Model for Repeated Measurements
MRI	Magnetic Resonance Imaging
MRI+	Positive for sacroiliitis (active inflammatory lesions)
MTX	Methotrexate
nr-axSpA	Non-radiographic axial SpondyloArthritis
NSAID	Non-Steroidal Anti-Inflammatory Drug
PCS	Physical Component Summary
PD	Pharmacodynamic
[redacted]	[redacted]
PRN	According to need, as required
PRO	Patient Reported Outcome
QoL	Quality of Life
RAN2	Extension Phase Randomized Set
SAE	Serious Adverse Event
SAF	Safety Set
SAF2	Extension Phase Safety Set
s.c.	Subcutaneous(ly)
SD	Standard Deviation
SE	Standard Error
SF-36	Medical Outcome Short Form (36) Health Survey
SPP	Safety Profiling Plan
TNF/TNF-α	Tumor Necrosis Factor
TNF-IR	TNF-α Inhibitor Incomplete Responders
ULN	Upper Limit of Normal
VAS	Visual Analog Scale

vs	Versus

1 Introduction

Data will be analyzed by Novartis according to the data analysis Section 9 of the study protocol. The statistical methodology for the final database lock of the AINH2315 trial is described below and any deviations from the protocol are documented. Additional detailed information regarding the analysis methodology is contained in the Appendix section.

2 Study objectives

There are two sets of primary and secondary objectives based on the regional regulatory precedent and feedback. These objectives were tested in separate analysis plans during the core phase and will not be repeated. For the purpose of the final database lock of the extension phase, only the exploratory objectives will be analyzed.

2.1 Primary objectives

Not applicable for the extension phase. Refer to [AIN457H2315 SAP] for details.

2.2 Secondary objectives

Not applicable for the extension phase. Refer to [AIN457H2315 SAP] for details.

3 Data presentation

The analysis of the extension phase data will be exploratory in nature and mainly reported by descriptive statistics, applying the same principles as in core phase data analysis.

Statistical models may be applied to produce confidence intervals for treatment summaries, but no formal statistical inference will be done.

Summary statistics for continuous variables will include N, mean, standard deviation (SD), minimum, lower quartile, median, upper quartile and maximum. Summary statistics for discrete variables will be presented in the number and percent of subjects in each category.

Efficacy data of the extension phase will be presented by the following 3 treatment groups:

- Core Phase Responders – Randomized AIN457 150 mg
- Core Phase Responders – Randomized AIN457 300 mg
- Core Phase Non-Responders – AIN457 300 mg Open-Label

Safety data will be summarized as follows for the entire treatment period depending on the timing of event:

- Any AIN457 150 mg
- Any AIN457 300 mg
- Any AIN457
- Placebo

Note that the treatment groups above for a subject may differ depending on the time period of the analysis.

Data will also be presented after patient up-titrated to secukinumab 300 mg open-label after Week 156, by a combination of the 'original' and 'current' treatment groups. This will represent the treatment combinations the subjects experience over the course of the entire trial in case of up-titration.

All listings will be presented by randomized treatments (Core Phase Responders) and Secukinumab 300 mg open-label but also indicating the current treatment.

4 Subjects and treatments

4.1 Analysis sets

The following analysis sets will be used in this trial:

Extension phase randomized set (RAN2): The RAN2 will be defined as all Core Phase Responders who were randomized. Unless otherwise specified, mis-randomized subjects (mis-randomized in IRT) will be excluded from the randomized set.

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

Subjects who have been randomized with an incorrect ASAS20 response status or prior to the availability of the ASAS20 criteria will be analyzed as randomized.

Extension phase full analysis set (EAS): The EAS will be comprised of all subjects enrolled in the extension phase (subjects from the randomized set plus the open label subjects) and study treatment has been assigned. Following the intent-to-treat principle, subjects will be evaluated according to the treatment assigned to at Week 104.

Safety set (SAF): The SAF includes all subjects who took at least one dose of study treatment during the treatment period in core study. Subjects will be evaluated according to treatment received.

Extension phase safety set (SAF2): The SAF2 includes all subjects enrolled in the extension phase and who took at least one dose of study treatment during the extension phase.

4.2 Treatment groups

Treatment groups for core phase responders are as follows:

- Secukinumab 150 mg

- Secukinumab 300 mg

and for core phase non-responders:

- Secukinumab 300 mg open-label

5 Subgroup definitions

Additional subgroup analysis may be conducted if deemed necessary.

6 Assessment windows, baseline and post baseline definitions, missing data handling

Baseline and post-baseline definitions

In general, a *baseline* value refers to the last measurement made prior to administration of the first dose of study treatment. A post-baseline value refers to a measurement taken after the first dose of study treatment.

Analysis visit windows

No analysis visit windows will be employed during the extension phase. Instead the nominal visits will be used. However, analysis visit window will be used for Week 104 to keep consistent with analyses in the core phase.

7 Subject disposition, background and demographic characteristics

7.1 Subject disposition

The number of subjects entered the extension will be presented. The number and percentage of subjects who completed the study periods and who discontinued the study prematurely (including the reason for discontinuation) will be presented at the end of the study, if appropriate, 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 for the entire study period. Protocol deviations due to COVID-19 will be reported in separate categories.

Additional analysis due to COVID-19 might be provided, e.g., summary of patients completed Week 120 prior to 1-Mar-2020 and patients missed visit and/or treatment in the extension phase due to COVID-19.

7.2 Background and demographic characteristics

The following common background and demographic variables will be summarized for patients entered the extension phase.

Continuous variables:

- Age (which is derived from date of birth and the date of informed consent)

- Height
- Weight
- Body mass index (BMI) = (body weight in kilograms) / (height in meters)^2

For BMI, height and body weight, the last value prior to randomization is used. If there is no weight recorded prior to taking of study drug, BMI will be missing.

Categorical variables:

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

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

- Patient's global assessment of disease activity and other ASAS components, hsCRP (mg/L and >ULN), ESR (mm/h), prior use (yes/no) of TNF-α inhibitors, use (yes/no) and separate dose of MTX (mg/week), sulfasalazine (g/day) and systemic corticosteroids (mg/day) at randomization, time since first diagnosis of nr-axSpA (years), time since onset of back pain, modified New York criteria for AS, HLA-B27, MASES, total back pain (VAS), nocturnal back pain (VAS), total BASDAI score, spinal pain (BASDAI question #2) and BASMI components (all seven in original units), presence of SIJ inflammation by MRI and each randomization strata level of the core phase (CRP+/MRI+, CRP+/MRI-, CRP-/MRI+).

Unless otherwise specified, summary statistics will be presented for continuous variables for each treatment group and for all subjects (total) in the EAS. 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 EAS.

8 Medical history

The evaluation of medical history was subject to prior interim analysis and will not be repeated. Refer to [AIN457H2315 SAP] for details.

9 Study treatment

The analysis of study treatment data will be based on SAF and SAF2. The number of active and placebo injections will be summarized by treatment group. 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.) during the extension phase will be presented.

Duration of exposure will be defined as the time from first dose of study 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) will have exposure to both medications 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.

10 Concomitant medication

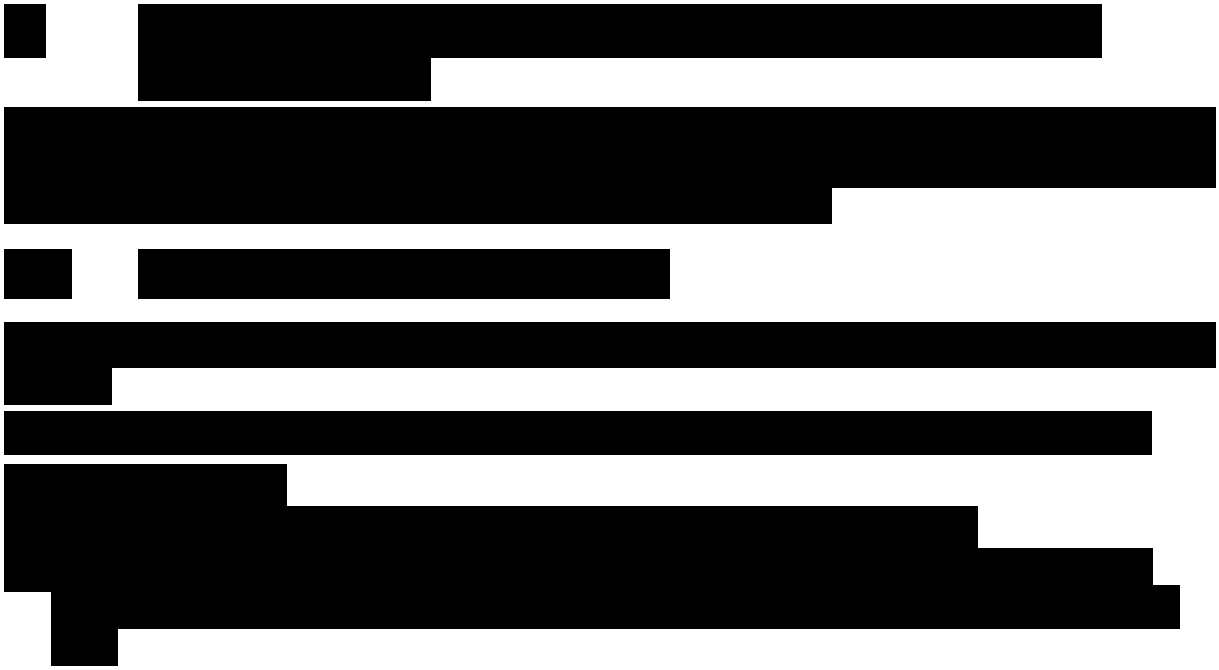
Concomitant medications will be summarized in separate tables by treatment group.

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

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



[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

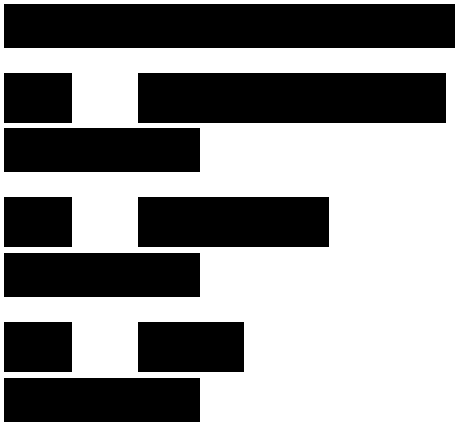
[redacted]

[redacted]

[redacted]

[redacted]

[redacted]



13 Safety evaluation

Summaries may be performed for the entire treatment period (including follow-up).

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.

13.1 Adverse events

The crude incidence of treatment emergent adverse events (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 date + 84 days) will be summarized by primary system organ class and preferred term. Confidence intervals for the crude rate will be derived. In addition, exposure time-adjusted incidence rates including 95% confidence intervals will be provided for the entire treatment period.

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, using a narrow search. The MedDRA version used for reporting the study will be described in a footnote.

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

An overview of the safety analyses which will be performed for treatment emergent AEs, on treatment labs and vital signs for each analysis period is described below.

Table 13-1 Overview of analyses on some safety endpoints

Analysis period	AEs & SAEs	AEs by severity	Study drug related AEs	AEs-SMQ	Risk	Notables for (vitals), lab criteria
Entire Treatment	• crude incidence • exposure time adjusted incidence	• crude incidence	• crude incidence	• exposure time adjusted incidence	• crude incidence • exposure time adjusted incidence	• crude 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 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
- at Level 1 for Risks and SMQ analyses

If adjudication is performed, the adjudication events (myocardial infarction, stroke, and cardiovascular death) will be listed.

Algorithms for date imputations will be provided in Programming Datasets Specifications.

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables on <on-treatment/treatment emergent> adverse events which are not serious adverse events with an frequency greater than 5% and on <on-treatment/treatment emergent> serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term on the SAF and SAF2 population.

13.2 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.

Reported laboratory assessments (hematology, chemistry and urinalysis) with either a less than or greater than sign (“<” or “>”) will be used for analysis after removal of the sign and conversion to standard unit. These laboratory data will be displayed in listings using the standard unit with the reported sign (“<” or “>”).

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:

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 13-2: 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.

Table 13-2 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	Life-threatening consequences; urgent intervention indicated
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; >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.

Newly occurring or worsening liver enzyme abnormalities will also be summarized based on the event criteria given below:

Table 13-3 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 laboratory criteria**) 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: 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.

13.3 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.

The number and percentage of subjects with newly occurring notable vital signs will be presented. The respective listing will be provided.

Criteria for notable vital sign abnormalities are provided below:

Table 13-4 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

13.4 Electrocardiogram (ECG)

Not applicable.

[redacted]

13.6 Compound specific safety evaluation

Safety topics of interest, such as risks defined in the Safety Profiling Plan (SPP), Risk Management Plan or topics of interest regarding signal detection or routine analysis are defined in the Program Case Retrieval Sheet that is stored in CREDI at the path Cabinets/CREDI Projects/A/AIN457A/Integrated Medical Safety.

The crude incidence and exposure-adjusted incidence rates for SPP risks will be summarized. In addition, listings will be provided presenting which subjects experienced which risk.

Important note: For the evaluation of SPP risks primary and secondary system organ classes of the MedDRA dictionary will be considered.

14 Determination of sample size

Please see [AIN457H2315 SAP] for the sample size calculation of the core phase.

The total sample size for extension phase is not pre-specified. It is estimated that approximately 70% of subjects enrolled in the core phase will enter the extension phase, which is equivalent to about 380 subjects. Around 305 subjects are estimated to be core phase responders and will be in the randomized group and around 75 subjects are expected to be core phase non-responders and will enter the open-label treatment group.

15 Appendix

15.1 Visit Windows

No analysis visit windows will be employed in the analysis, except for Week 104. Refer to [AIN457H2315 SAP_Week104 Amendment 1] for details about the analysis visit window of Week 104.

All data collected will be displayed in listings.

15.2 Statistical methodology and assumptions

15.2.1 Analysis of continuous data

15.2.1.1 Summary statistics for continuous data

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

15.2.1.2 Analysis of covariance

Endpoints with continuous data type expected to be normally distributed (e.g., change from baseline in BASDAI) will be analyzed using a mixed-effects repeated measures model (MMRM) with factors, covariates, interactions and covariance structure as specified for respective analysis. The SAS code below outlines a template for the analysis where covariates and interaction terms can be added or removed as required:

ods output diffs=*lsm_diff* lsmeans=*lsm_est*;

proc mixed data=*mydata*;

class *treatment strata tnf_status visit*;

model *outcome* = *treatment strata tnf_status visit baseline_score baseline_weight*

*treatment*visit baseline_score *visit* / ddfm=kr;

repeated *visit* / type=un subject=;

lsmeans *treatment*visit* / diff cl;

run;

Least-square-mean (LSM) estimates and the standard error (SE) for each treatment groupwill be calculated at appropriate analysis visits, i.e., Week 104, Week 112 and Week 120.

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

1. change to ddfm=bw. If still no convergence, perform step 2.
2. change to type=cs. If still no convergence, perform step 3.
3. remove covariates in the following order until convergence: *baseline_weight*, *baseline_score*visit*, *tnf_status*, *strata*.

15.2.2 Analysis of binary and categorical data

15.2.2.1 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-α/2), n as total number of subjects (i.e., number of subjects in the denominator), p as estimated crude incidence (number of subjects with event / n) and $q = 1-p$

Then the lower limit is

$$L = 100 \times \max\left(0, \frac{2np + z^2 - 1 - z\sqrt{z^2 - 2 - 1/n + 4p(nq+1)}}{2(n+z^2)}\right)$$

and the upper limit is

$$U = 100 \times \min\left(1, \frac{2np + z^2 + 1 + z\sqrt{z^2 + 2 - 1/n + 4p(nq-1)}}{2(n+z^2)}\right).$$

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

For binary response variables the placebo-adjusted response rates including 95% confidence interval will be derived.

SAS code for risk difference:

```
proc freq data=;
    tables response* treatment / riskdiff;
run;
```

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

15.2.2.2 Multiple Imputation

A linear regression model will be used to perform multiple imputation (MI) under a missing-at-random (MAR) assumption. To help preserve the relationship between outcome and covariates within each treatment a separate model will be run for each treatment. This will also help ensure that the imputation model does not make stronger assumptions on data relations than the analysis model.

The SAS code below outlines a template for the analysis where covariates and visits can be added or removed as required. To ensure that results can be replicated the data should be sorted by treatment group and unique subject ID before running the model (the data should be in horizontal format with one subject per dataset row).

```
proc mi data=mydata seed=4572315 nimpute=100 out=mi_out;
    by treatment;
    class strata tnf_status;
    fcs reg (/details);
    var strata tnf_status baseline_weight value1 value2 value3 value4;
run;
```

In the template code, the continuous variable to be imputed is *value* (e.g., *value1* could be the baseline value and *value2* the first post-treatment measurement of the variable to be imputed.) Normally, all data collection visits during the analysis period of interest (i.e., Week 104, Week 112 and Week 120) would be included in the model. Including variables using a CLASS statement instead of a BY statement should help facilitate model convergence also when the number of non-missing data points are low for some specific covariate level and visit combination. The FCS option is used to ensure that also non-monotone missing data can be handled in an appropriate way.

If convergence is not obtained the following sequential steps will be used:

1. remove *baseline_weight*. If still no convergence, perform step 2.
2. remove *tnf_status*. If still no convergence, perform step 3.
3. remove *status*.

For a situation where several variables need to be imputed using separate models (e.g., using independent models to impute each component needed to derive a response variable *V*) a step-wise process needs to be implemented as outlined below:

1. Run the SAS code as described above for the first variable to be imputed
2. Run the SAS code as described above for the next variable to be imputed (but with the following changes: "data=*mi_out*", "out=*mi_out2*", "by *treatment _imputation_*", "nimpute=1")
3. Repeat step 2, but with input dataset equal to the output dataset from the prior step, until all *j* variables have been imputed resulting in a dataset named mi_out_j
4. Derive the variable *V* from within mi_out_j

To obtain binary response rates and confidence intervals for individual treatment groups the following process should be followed (exemplified for one visit):

```
ods output binomialprop=bin_est;

proc freq data=;
    by treatment _imputation_;
    table response / binomial (cl=wilson correct);
run;
```

Then apply a logit transformation on the saved proportions and derive its standard error:

```
data bin_est;
    set bin_est;
    estimate=log(_bin_/(1-_bin_));
    estimate_se=e_bin/(_bin_*(1-_bin_));
run;
```

The transformed binomial proportion estimates and its standard errors are then combined by applying Rubin's rules as described above using PROC MIANALYZE. Before presenting the combined data it needs to be transformed back as follows:

data *mi_result*;

set *mi_result*;

prop_est=1/(1+exp(-*estimate*));

prop_lower=1/(1+exp(-*lclmean*));

prop_upper=1/(1+exp(-*uclmean*));

run;

If all responses are imputed as 0 (or 1) for all imputation datasets for a specific treatment group then the between-imputation-variation will be zero. The combined final response rate would be presented as seen in any of the imputed datasets but the 95% CI will be undefined.

If after imputation all responses are either 0 or 1 for a combination of treatment group and imputation dataset it will not be possible to perform a logit transformation and the response rate (0% or 100%) will be presented without 95% CI.

15.2.3 Crude incidence and related risk estimates

15.2.3.1 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 $(1-\alpha/2)$-quantile of the standard normal distribution (SAS: z = PROBIT(1-α/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.

15.2.4 Exposure adjusted incidence rate and related risk estimates

15.2.4.1 Exposure adjusted incidence rate and 100*(1-α)% confidence interval

It will be assumed that for each of n subjects in a clinical trial the time t_j (j=1,...,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 θ. The rate parameter θ 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-α)% confidence interval for a Poisson variable with parameter θT and observed value D can be obtained based on (Garwood 1936), from which an exact 100*(1-α)% confidence interval for D/T will be derived as follows (Sahai 1993; Ulm 1990):

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

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

Where $c_{\alpha,k}$ is the αth 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 15-1 Examples for calculating exposure time for incidence rates (IR)

1st treatment / total exposure time	2nd treatment / total exposure time	AE event onset (in days from study start)	Exposure for IR
Placebo / 100 days	AIN457 150 mg / 200 days	Day 50 (during 1st treatment) Day 110 (10 days into 2nd treatment)	Placebo: 50 days AIN457 150 mg: 10 days Any AIN: 10 days

16 References

Braun J, McHugh N, Singh A, Wajdula JS, Sato R (2009) Improvement in patient-reported outcomes for patients with ankylosing spondylitis treated with etanercept 50 mg once-weekly and 25 mg twice-weekly. Rheumatology (Oxford); 46: 999-1004

Cella DF, Tulsky DS, Gray G et al. (1993) The Functional Assessment of Cancer Therapy (FACT) scale: Development and validation of the general measure. Journal of Clinical Oncology; 11(3):570-579.

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

Dougados M, Paternotte S, et al. (2010) ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studied in axial spondyloarthritis. Ann Rheum Dis; 70:249-251

Doward LC, Spoorenberg A, Cook SA, et al. (2003) Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. Ann Rheum Dis; 62(1):20-26.

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

Haywood KL, et al. (2002) Disease-specific, patient-assessed measures of health outcome in ankylosing spondylitis: reliability, validity and responsiveness. Rheumatology; 41:1295-1302.

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

Lukas C, Landewé R, Sieper J. (2009) Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis.; 68(1):18-24.

Machado PM, Landewé RB, van der Heijde DM. (2011) Endorsement of definitions of disease activity states and improvement scores for the ankylosing spondylitis Disease Activity Score: results from OMERACT 10. J Rheum; 38(7):1502-1506.

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

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

Sieper J, Rudwaleit M, Baraliakos X. (2009) The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis.; 68 Suppl 2:ii1-44.

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 D, Dijkmans B, et al. (2005) Efficacy and safety of infliximab in patients with ankylosing spondylitis: Results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum; 52(2):582-591.

van Tubergen A, Landewe R, van der Heijde D, et al. (2001) Combined spa-exercise therapy is effective in patients with ankylosing spondylitis: a randomized controlled trial. Arthritis Rheum; 45: 430–38