NOVARTIS

Clinical Development

Secukinumab (AIN457)

CAIN457P12301 / NCT04156620

A randomized, double-blind, placebo-controlled, parallel group, phase III multicenter study of intravenous secukinumab to compare efficacy at 16 weeks with placebo and to assess safety and tolerability up to 52 weeks in subjects with active Ankylosing Spondylitis or non-radiographic axial SpondyloArthritistudy

Statistical Analysis Plan (SAP)

Author: [redacted], Trial Statistician

Document type: SAP Documentation

Document status: Final

Release date: 12-Jan-2023

Number of pages: 45

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
04-Nov-2022	Initial version	Creation of final version This SAP is created for the Final DBL. Analyses that are specific for Week 16 DBL and will not be repeated in the Final DBL are removed, e.g., inferential analyses, baseline information, etc. [redacted]	N/A - First version	NA
05-Dec-2022	Ammendment 1	[redacted] Duration of exposure updated for patients who discontinue treatment, to be consistent within the program.		

Table of contents

List of tables

List of abbreviations

AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANCOVA	Analysis of Covariance
AS	Ankylosing spondylitis
ASAS	Assessment of SpondyloArthritis International Society
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASQoL	Ankylosing Spondylitis Quality of Life
AST	Aspartate Aminotransferase
axSpA	Axial spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
[redacted]	[redacted]
BMI	Body Mass Index
BSL	Baseline
CRF	Case Report/Record Form (paper or electronic)
CRP	C-Reactive Protein
CSR	Clincial Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DMARD	Disease Modifying Anti-rheumatic Drug
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
FAS	Full Analysis Set
FDA	Food and Drug Administration
GGT	Gamma Glutamyl Transferase
H	Hour
HDL	High Density Lipoprotein
HLA	Human Leukocyte Antigen
hsCRP	high sensitivity C-Reactive Protein
i.v.	Intravenous
[redacted]	[redacted]
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
kg	kilogram(s)
MAR	Missing At Random
[redacted]	[redacted]

MCS	Mental Component Summary
MedDRA	Medical Dictionary for Regulatory Activities
Mg	milligram(s)
MI	Multiple Imputation
MMRM	Mixed effect Model for Repeated Measurements
MTX	Methotrexate
nr-axSpA	non-radiographic axial spondyloarthritis
NSAID	Non-Steroidal Anti-Inflammatory Drug
PCS	Physical Component Summary
PD	Pharmacodynamic(s)
PK	Pharmacokinetic(s)
PRN	as needed
PRO	Patient Reported Outcome
PSQI	Pittsburgh Sleep Quality Index
PT	Preferred term
q4w	Every 4 weeks
QoL	Quality of Life
RMP	Risk Management Plan
SAE	Serious Adverse Event
SD	Standard Deviation
SF-36	Medical Outcome Short Form (36) Health Survey
SF-36 PCS	Short Form-36 Physical Component Summary
SOC	System Organ Class
SMQ	Standardized MedDRA Query
SpA	Spondyloarthritis
[redacted]	[redacted]
TBL	Total Bilirubin
TNF-IR	TNFα Inhibitor Incomplete Responder
TNF/TNFα	Tumor Necrosis Factor
ULN	Upper Limit of Normal
VAS	Visual Analog Scale
vs	Versus
[redacted]	[redacted]

1 Introduction

Data will be analyzed by Novartis according to the data analysis Section 12 of the clinical study protocol. The statistical methodology is described below and any deviations from the protocol are documented. Additional detailed information regarding the analysis methodology is contained in the Appendix section.

1.1 Study design

This multicenter study uses a randomized, double-blind, placebo-controlled, parallel-group design to study the efficacy, safety, and tolerability of treatment with intravenous secukinumab (initial dose of 6 mg/kg followed thereafter with 3 mg/kg administered every four weeks starting at Week 4) in subjects with active axSpA. The study population consists of approximately 400 subjects with active AS and approximately 100 subjects with active nr-axSpA, despite current or previous NSAID, conventional DMARD and / or TNF inhibitor therapy, or intolerance to these therapies. A screening (SCR) period of up to 10 weeks will be used to assess eligibility, followed by randomization to 52 weeks of study treatment.

At baseline, subjects with active AS and nr-axSpA will be randomized to one of the two treatment groups:

- Group 1: approximately 200 AS subjects and approximately 50 nr-axSpA subjects; These subjects will receive secukinumab 6 mg/kg i.v. at randomization (Baseline (BSL) visit), followed by the administration of secukinumab 3 mg/kg i.v. every four weeks starting at Week 4 through Week 48 (exposure through Week 52).
- Group 2: approximately 200 AS subjects and approximately 50 nr-axSpA subjects; These subjects will receive i.v. placebo at randomization (BSL visit), Weeks 4, 8, and 12 , followed by the administration of secukinumab 3 mg/kg i.v. at Week 16 and every four weeks through Week 48 (exposure through Week 52).

This study will consist of 4 periods totaling 70 weeks: the screening period (up to 10 weeks), the treatment period 1 (total duration of 16 weeks) and the treatment period 2 (total duration of 36 weeks) followed by the safety follow up period of 8 weeks after the end of treatment visit (i.e., Week 52).

The subjects will be stratified at randomization according to disease condition (i.e., AS or nr-axSpA). No more than 20% TNF Inhibitor Incomplete Responders (TNF-IR) subjects will be enrolled in the study. Starting at Week 16, all subjects will switch to open-label intravenous secukinumab, including all placebo subjects.

No subject will be on placebo treatment after Week 16, because axSpA is considered a chronic disease with no ‘true’ and lasting placebo response. However, all subjects and investigators/site staff will remain blinded to the original randomized treatment group assignment (secukinumab treatment vs placebo). Study treatment will continue up to Week 52. An end of treatment visit (i.e., Week 52) is to be done 4 weeks after last study treatment administration and a post treatment follow-up visit (i.e., Week 60) is to be done 8 weeks after the end of treatment visit for all subjects (regardless of whether they complete the entire study as planned or discontinue prematurely). All i.v. infusions will be performed at the study site and site personnel will administer the infusions to subjects.

Rescue medication is not allowed until Week 16. However, subjects who are deemed by the investigator not to be benefiting from the study treatment based on safety and efficacy assessments or for any reason of their own accord will be free to discontinue participation in the study at any time. The study will have a primary endpoint analysis at Week 16. Therefore, the primary analysis will be performed with Week 16 data once the last subject has completed the Treatment Period 1.

The primary endpoint analysis will be performed after all subjects complete Week 16 in order to support regulatory filing. As the primary analysis is scheduled at Week 16, no adjustment will be made to the testing strategy to control family-wise type I error rate for this analysis.

1.2 Study objectives and endpoints

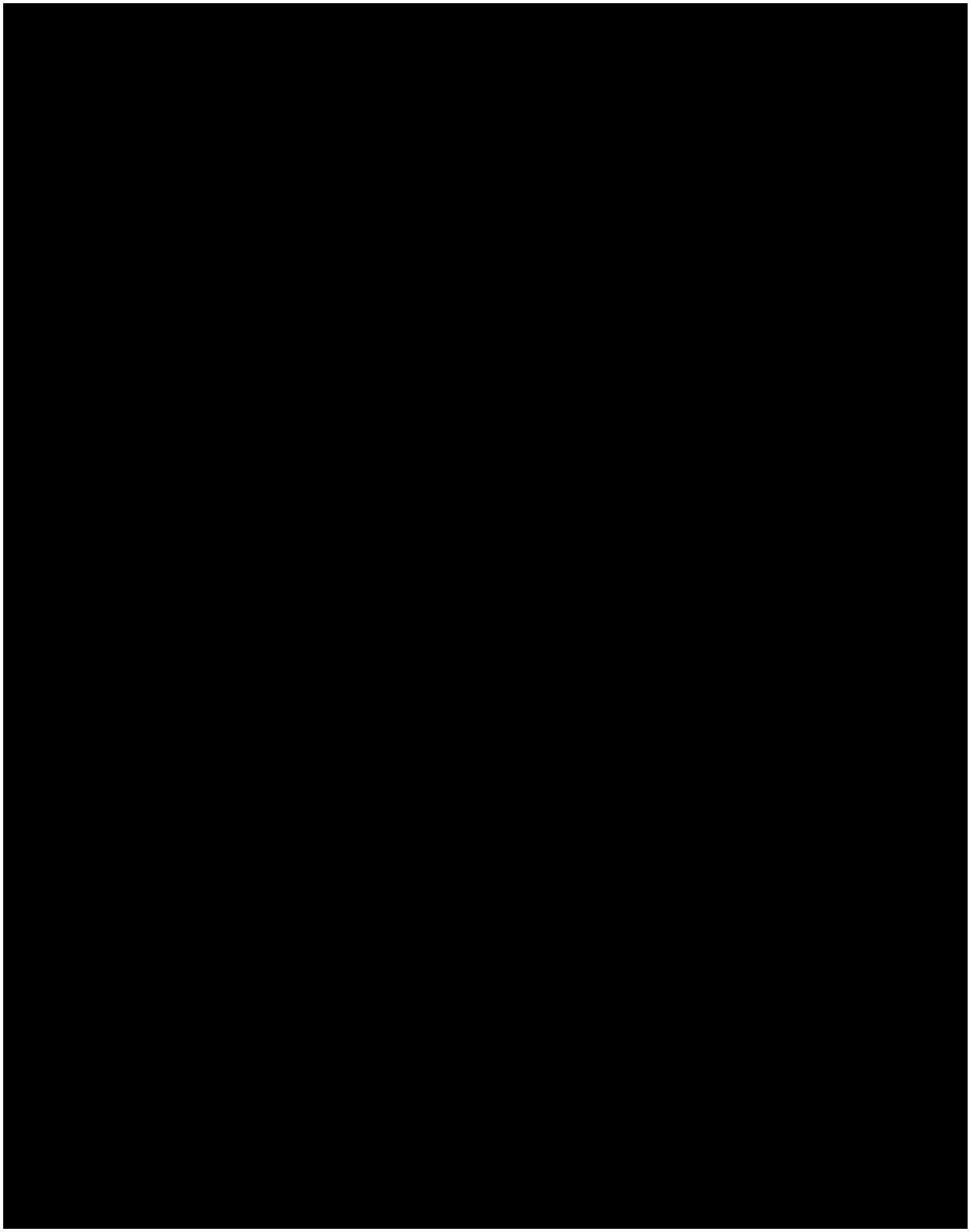
1.2.1 Primary objective

The primary objective is to demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo in subjects with active axSpA (AS and nr-axSpA) based on the proportion of subjects achieving an ASAS40 (Assessment of SpondyloArthritis International Society criteria) response.

1.2.2 Secondary objectives

1. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the proportion of subjects achieving Ankylosing Spondylitis Disease Activity Score (ASDAS)-C-Reactive Protein (CRP) major improvement
2. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the change from baseline in total Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
3. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the proportion of subjects meeting the ASAS 5/6 response criteria
4. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the change from baseline of Bath Ankylosing Spondylitis Functional Index (BASFI)
5. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the change from baseline in Short Form-36 Physical Component Summary (SF-36 PCS)
6. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the change from baseline in Ankylosing Spondylitis Quality of Life (ASQoL) scores
7. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the change from baseline in high sensitivity C-Reactive Protein (hsCRP)
8. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the proportion of subjects achieving an ASAS20 response
9. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the proportion of subjects achieving Ankylosing Spondylitis Disease Activity Score (ASDAS)-C-Reactive Protein (CRP) inactive disease
10. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the proportion of subjects achieving ASAS partial remission

11. To demonstrate that the efficacy of i.v. secukinumab at Week 16 is superior to placebo based on the change from baseline in Pittsburgh Sleep Quality Index (PSQI)
12. Overall safety and tolerability of i.v. secukinumab compared to placebo as assessed by physical exams, vital signs, laboratory assessments and adverse event monitoring



[redacted]

2 Statistical methods

2.1 Data analysis general information

To support the submission, tables and figures will be generated for overall axSpA population. Listings will be generated for overall axSpA population and will be sorted by disease condition (AS and nr-axSpA), except for the listing of randomization allocation to treatment which will be sorted by the randomization number.

Summary statistics for continuous variables will include N, mean, standard deviation, 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.

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 2 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). The 2 treatment groups represent the regimens subjects will be eligible to be randomized to for the first 16 weeks of the study.

- Secukinumab regimen: approximately 200 AS subjects and approximately 50 nr-axSpA subjects; These subjects will receive secukinumab 6 mg/kg i.v. at randomization (Baseline (BSL) visit), followed by the administration of secukinumab 3 mg/kg i.v. every four weeks starting at Week 4 through Week 48 (exposure through Week 52).
- Placebo regimen: approximately 200 AS subjects and approximately 50 nr-axSpA subjects; These subjects will receive i.v. placebo at randomization (BSL visit), Weeks 4, 8, and 12 , followed by the administration of secukinumab 3 mg/kg i.v. at Week 16 and every four weeks through Week 48 (exposure through Week 52).

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.

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 AIN457P12301, this is the first 16-weeks of treatment. Comparative efficacy will be performed based on the Full Analysis Set (FAS) population using the randomized treatment. Data collected after Week 16 will generally be summarized descriptively on the FAS population using treatment sequence.

Efficacy data following treatment switch

Data will also be presented after Week 16, by a combination of the ‘original’ and ‘switch’ treatment groups and will be referred to as treatment sequence. These treatment sequences represent the treatment combinations the subjects experience over the course of the entire trial

in case of switch (e.g., placebo patients who are reassigned to secukinumab 3 mg/kg i.v. at Week 16).

All listings will be presented by treatment sequence.

2.2 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.

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

2.2.1 Treatment groups

The summaries by treatment will 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 6 mg/kg - 3 mg/kg i.v.
  - Placebo
- Treatment sequence:
  - AIN457 6 mg/kg - 3 mg/kg i.v.
  - Placebo - AIN457 3 mg/kg i.v.
  - Placebo Not Switch
- Switch treatments (for placebo patients who cross-over):
  - AIN457 3 mg/kg i.v. (will only be counted under Any AIN457)

2.2.2 Subgroup of interest

The primary and secondary endpoints will be analyzed by disease condition (AS and nr-axSpA). Analyses by previous use of TNF-α inhibitor may also be provided for overall axSpA population.

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

The number of subjects screened will be presented. In addition, the reasons for screen failures will be provided. The number and percentage of subjects in the randomized set who completed the study periods and who discontinued the study prematurely (including the reason for discontinuation) will be presented at the end of each treatment period (Week 16 and Week 52) and follow-up period, 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. 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 participants missed visit and/or treatment due to COVID-19.

2.3.2 Background and demographic characteristics

Data on background and demographic characteristics have been summarized in the Week 16 Clincial Study Report (CSR) and will not be reproduced for the Final Analysis.

2.3.3 Medical history

Data on medical history and smoking history have been summarized in Week 16 CSR and will not be reproduced for the Final Analysis.

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

2.4.1 Study treatment / compliance

The analysis of study treatment data will be based on the safety set. The number of active and placebo i.v. infusions received will be presented 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, ≥ 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 from placebo to open-label secukinumab) 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 open-label secukinumab treatment at Week 16) 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.

2.4.2 Prior, concomitant and post therapies

Prior medications have been summarized in the Week 16 CSR and will not be reproduced for the Final Analysis.

Any medication given at least once between the day of first dose of randomized 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.

Concomitant medications will be presented in alphabetical order, by Anatomical Therapeutic Classification (ATC) codes and grouped by anatomical main group (the 1st level of the ATC codes). 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 non-drug therapies and procedures will be summarized by primary system organ class and MedDRA preferred term.

Any non-drug therapies and procedures done between the day of first dose of study treatment and within 84 days after last dose will be defined as concomitant non-drug therapies and procedures, including those which were started pre-baseline and continued into the period where study treatment is administered.

The number and percentage of subjects receiving concomitant axSpA therapy will be presented by randomized treatment group and the total duration of previous exposure to axSpA therapies.

[redacted]

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

2.5 Analysis of the primary objective

The primary efficacy variable is response to treatment according to the ASAS40 criteria at Week 16 in patients with active axSpA (AS and nr-axSpA).

The analysis of primary variable at Week 16 will not be reproduced for the Final Analysis.

2.6 Analysis of secondary objectives

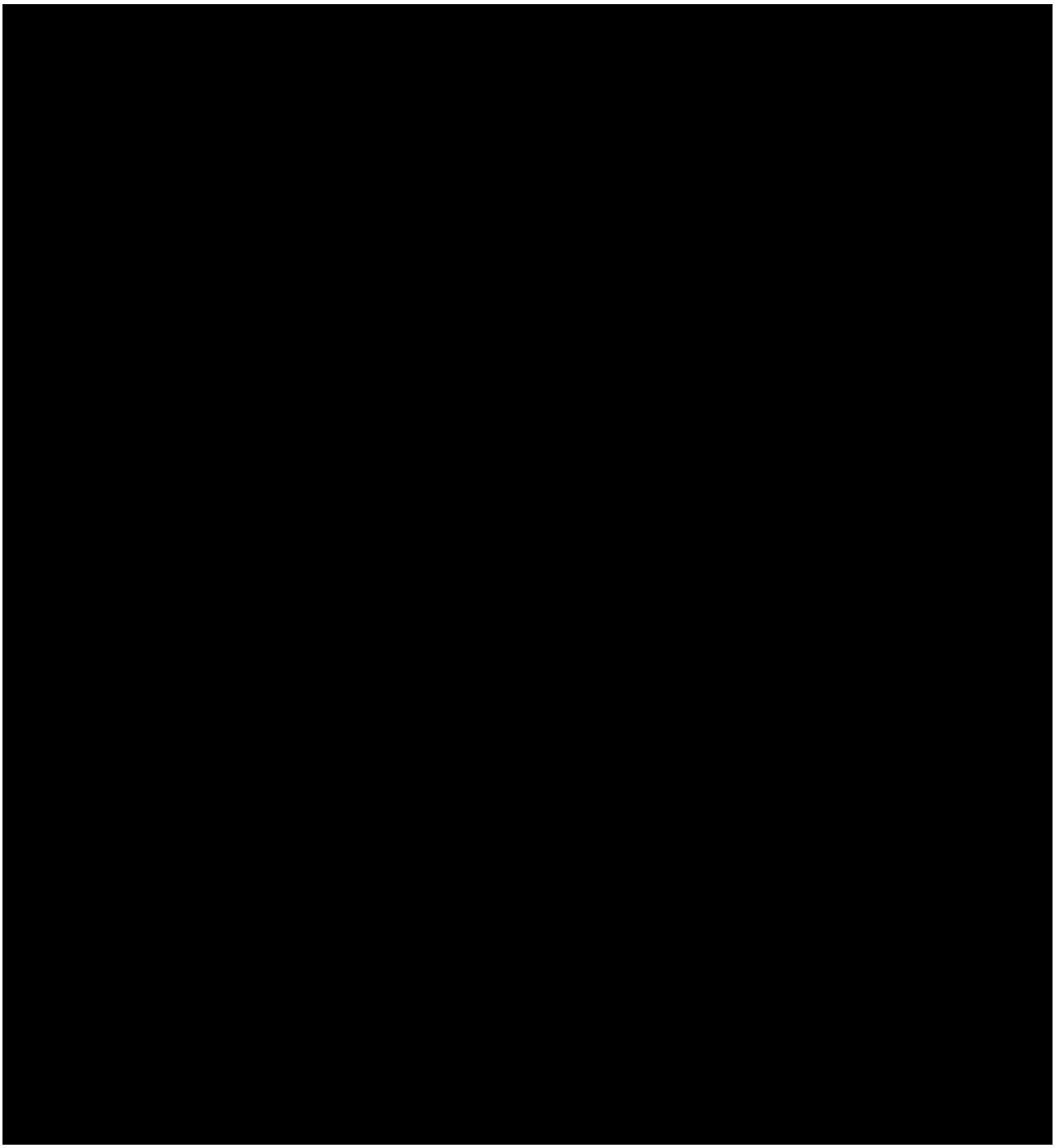
The secondary efficacy variables include,

- response to treatment at Week 16 according to the ASDAS-CRP major improvement criteria
- change from baseline in BASDAI at Week 16
- response to treatment at Week 16 according to the ASAS 5/6 criteria
- change from baseline in total BASFI score at Week 16
- change from baseline in SF-36 PCS at Week 16
- change from baseline in ASQoL at Week 16
- change from baseline in hsCRP at Week 16
- response to treatment at Week 16 according to the ASAS20 criteria

- response to treatment at Week 16 according to the ASDAS-CRP inactive disease criteria
- response to treatment at Week 16 according to the ASAS partial remission
- change from baseline in PSQI at Week 16

The analysis of secondary variables at Week 16 will not be reproduced for the Final Analysis.

[redacted]



2.8 Safety analyses

Summaries may performed separately for the entire treatment period. The analyses of the follow-up period will be limited to summaries for treatment-emergent adverse events, serious adverse events and risks based on adverse events.

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.

In addition for subjects who discontinue study treatment but continue with study participation, an additional AE listing will be prepared displaying which events occurred after the study treatment discontinuation.

2.8.1 Adverse events (AEs)

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 as described in Section 5.4.3. In addition, exposure time-adjusted incidence rates including 95% confidence intervals will be provided for the entire treatment period (see Section 5.4.4). 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 the secukinumab group 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 requiring concomitant medication.

Adverse events will also be reported separately by Standardized MedDRA Query (SMQ) according to MedDRA, using a narrow search. The MedDRA version used for reporting the study will be described in a footnote.

A listing of non-treatment emergent adverse events will be provided. These adverse events occurred before the first dose of the study treatment. The crude incidence rate will be provided without treatment information.

For SAEs occurred during screening a listing will be prepared for all subjects screened including screening failures.

An overview of the safety analyses which will be performed for treatment emergent AEs, labs and vital signs for each analysis period is described in Table 2-1.

Table 2-1 Overview of analyses on some safety endpoints

Analysis period	AEs & SAEs	AEs by severity	Study drug related AEs	AEs-SMQ	Risk	Notables for (vitals), lab criteria
Day 1 – Week 16	• 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:

- Primary SOC level for AE and SAE
- Level 1 for Risks and SMQ analyses
- PT level for AE and SAE

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 incidence greater than *x*% 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 safety set population. Here threshold value *x* is set to 2 (%), but it may be updated following review of the dry run outputs.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE/SAE suspected to be related to study treatment/non SAE has to be checked in a block e.g., among AEs in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

2.8.2 Laboratory data

The summary of lab data will only include treatment emergent 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, given in Table 2-2: hemoglobin, platelets, white blood cell count, neutrophils, lymphocytes, creatinine, total bilirubin (TBL), gamma-glutamyl transferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), glucose, cholesterol, triglycerides (TG).

These summaries will be split into hematology and chemistry for study level reports and the pooled summary of clinical safety.

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

CTCAE v4.03 Term	Grade 1	Grade 2	Grade 3	Grade 4
HGB decreased (Anemia)	<LLN – 100 g/L	<100 – 80 g/L	<80 g/L	See note below
Platelet count decreased	<LLN – 75.0 × 10e9 /L	<75.0 - 50.0 × 10e9 /L	<50.0 – 25.0 × 10e9 /L	<25.0 × 10e9 /L
White blood cell decreased	<LLN - 3.0 × 10e9 /L	<3.0 - 2.0 × 10e9 /L	<2.0 - 1.0 × 10e9 /L	<1.0 × 10e9 /L
Neutrophil count decreased	<LLN - 1.5 × 10e9 /L	<1.5 - 1.0 × 10e9 /L	<1.0 - 0.5 × 10e9 /L	<0.5 × 10e9 /L
Lymphocyte count decreased	<LLN - 0.8 × 10e9/L	<0.8 - 0.5 × 10e9 /L	<0.5 - 0.2 × 10e9 /L	<0.2 × 10e9 /L
Creatinine increased*	>1 - 1.5 × baseline; >ULN - 1.5 × ULN	>1.5 - 3.0 × baseline; >1.5 - 3.0 x ULN	>3.0 baseline; >3.0 - 6.0 × ULN	>6.0 × ULN
TBL increased	>ULN - 1.5 × ULN	>1.5 - 3.0 × ULN	>3.0 - 10.0 × ULN	>10.0 × ULN
GGT increased	>ULN - 2.5 × ULN	>2.5 - 5.0 × ULN	>5.0 - 20.0 × ULN	>20.0 × ULN
ALT increased	>ULN - 3.0 × ULN	>3.0 - 5.0 × ULN	>5.0 - 20.0 × ULN	>20.0 × ULN
AST increased	>ULN - 3.0 × ULN	>3.0 - 5.0 × ULN	>5.0 - 20.0 × ULN	>20.0 × ULN
ALP increased	>ULN - 2.5 × ULN	>2.5 - 5.0 × ULN	>5.0 - 20.0 × ULN	>20.0 × ULN
Glucose increased (Hyperglycemia)	>ULN - 8.9 mmol/L	>8.9 - 13.9 mmol/L	>13.9 - 27.8 mmol/L	>27.8 mmol/L

CTCAE v4.03 Term	Grade 1	Grade 2	Grade 3	Grade 4
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: Grade 4 Hemoglobin events are defined as life-threatening anemia events and will not be displayed in the table, as a numerical range is not provided in the CTCAE.

*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 × LLN
- LDL, cholesterol, triglycerides:
  - >=ULN
  - >1.5 × ULN
  - >2.5 × ULN

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

Table 2-3 Liver-related events

Parameter	Criterion
ALT	>3×ULN; >5×ULN; >8×ULN;>10×ULN, >20×ULN
AST	>3×ULN; >5×ULN; >8×ULN >10×ULN; >20×ULN
ALT or AST	>3×ULN; >5×ULN; >8×ULN >10×ULN; >20×ULN
TBL	>1.5×ULN, >2×ULN, >3×ULN,
ALP	>2×ULN, >3×ULN. >5×ULN
ALT or AST & TBL	ALT or AST>3×ULN & TBL >2×ULN; ALT or AST >5×ULN & TBL >2×ULN; ALT or AST >8×ULN & TBL >2×ULN; ALT or AST >10×ULN & TBL >2×ULN
ALP & TBL	ALP >3×ULN & TBL >2×ULN ALP >5×ULN & TBL >2×ULN
ALT or AST & TBL & ALP	ALT or AST>3×ULN & TBL >2×ULN & ALP <2×ULN **(Hy's Law**) Note: elevated ALP may suggest obstruction as a consequence of gall bladder or bile duct disease; ALP may also be increased in malignancy. FDA therefore terms Hy's Law cases as indicators of *pure hepatocellular injury*. This does not mean that cases of ALT or AST >3×ULN & TBL >2×ULN & ALP ≥2×ULN 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 = 6.42×ULN is counted for ALT >3×ULN and ALT>5×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).

Individual subject data listing will be provided for hematology.

2.8.3 Other safety data

2.8.3.1 Vital signs

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

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

change from baseline = post-baseline value – baseline value

The number and percentage of subjects with newly occurring notable vital signs will be presented. Criteria for notable vital sign abnormalities are provided in Table 2-4:

Table 2-4 Criteria for notable vital sign abnormalities

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

[redacted]

2.8.3.3 Compound specific safety evaluation

Safety topics of interest, such as risks defined in the Risk Management Plan (RMP) 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 RMP risks will be summarized. In addition, listings will be provided presenting which subjects experienced which risk.

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

2.9 Pharmacokinetic endpoints

During modeling of the PK of secukinumab, the broad principles outlined in the [FDA Guidance for Industry: Population Pharmacokinetics] will be followed.

All completed subjects with quantifiable pharmacokinetic (PK) measurements will be included in the PK data analysis. Serum concentrations will be expressed in mass per volume units. All concentrations below the limit of quantification as well as missing data will be labeled as such in the concentration data listings. PK concentrations will be summarized by visit and treatment group. In addition to mean, standard deviation (SD), coefficient of variation (CV), median and quartiles, the geometric mean and geometric coefficient of variation (CV), minimum and maximum, as well as n and n(log) will be presented. The formula for deriving the geometric mean and CV (%) is as follows:

- CV (%) = (SD/mean)*100,
- geometric mean=exp((sum of log transformed data) / number of non-missing data points after log transformation),
- geometric CV = sqrt(exp(variance of log-transformed data)-1)*100.

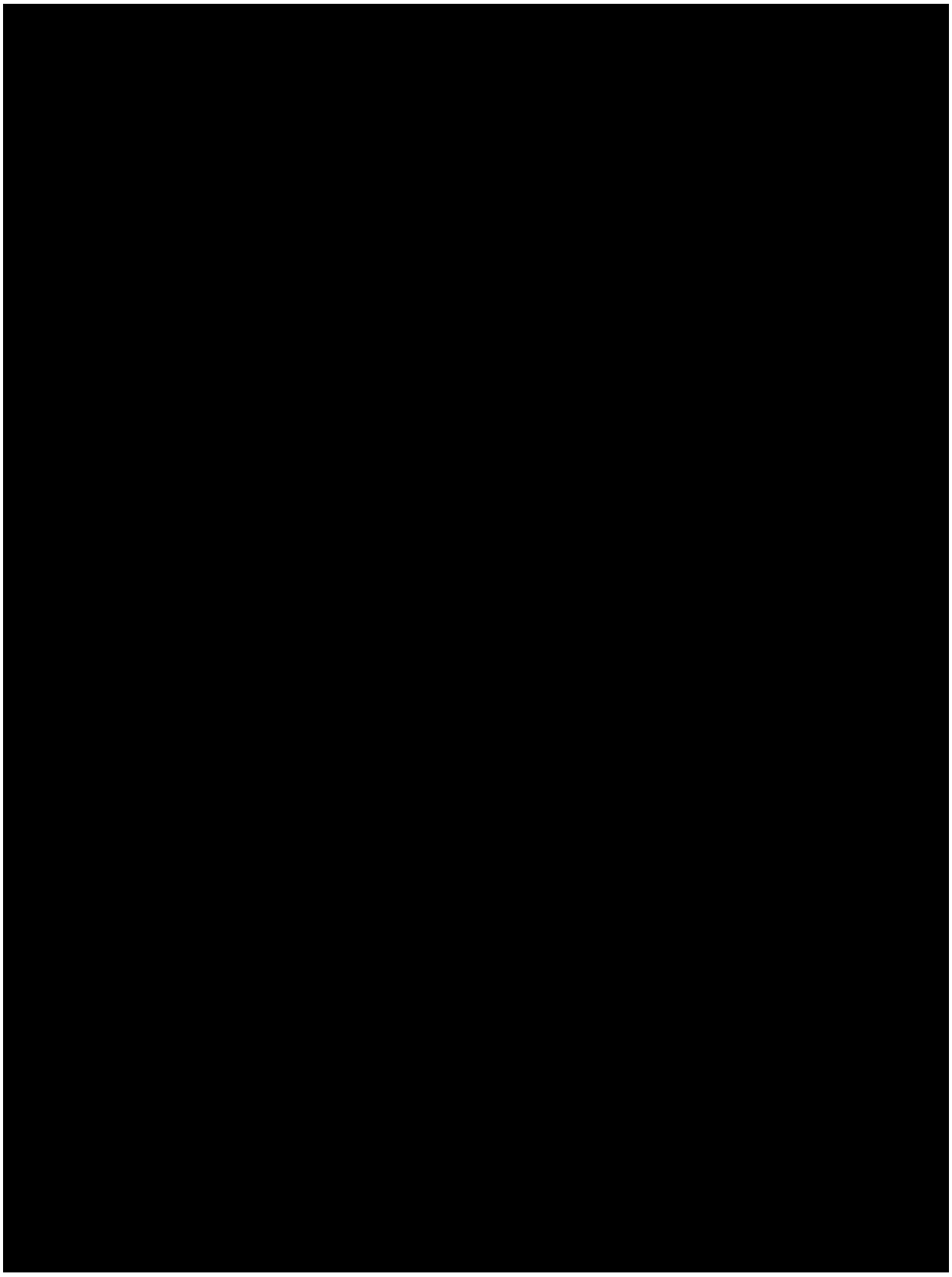
In addition, sample number, concentration, sample date, sample time and elapsed time from first dose (in days) will be listed by treatment sequence.

Values below lower limit of quantification/below detection limit will be imputed by 0.

Pharmacokinetic data of the study treatment will be analyzed with a population-pharmacokinetic mixed effects model. The analysis will be based on a pooled data set, including PK samples from previous studies. The modeling approach will be further detailed in a modeling plan. Results will be reported separately.

2.10 PD [redacted] analyses

[redacted]



2.11 Patient reported outcomes (PROs)

Patient reported outcomes will be evaluated based on FAS unless otherwise specified.

BASDAI

The following variables will be evaluated:

- BASDAI
- [redacted]

For BASDAI, summary statistics of observed data by visit and change from baseline in BASDAI will be provided for each treatment.

[redacted]

BASFI

See Section 5.1 Description of efficacy variables for details.

Patient's global assessment of disease activity (VAS)

Summary statistics of observed data by visit and change from baseline in patient's global assessment of disease activity will be provided for each treatment.

Patient's assessment of back pain intensity (VAS)

Summary statistics of observed data by visit and change from baseline in patient's assessment of back pain intensity will be provided for each treatment.

SF-36

The following variables will be evaluated:

- SF-36 domain scores (based on a scale of 0-100)
- SF-36 PCS [redacted] (norm-based scores)
- SF-36 PCS [redacted] (improvement of ≥ 2.5 points, Lubeck 2004)

For the change from baseline in SF-36 summary scores (PCS [redacted]), summary statistics will be provided using observed data for each treatment regimen.

For the responder analyses, the proportion of responders will be descriptively summarized along with its 95% CI for each randomized treatment based on observed data.

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

ASQoL

Summary statistics of observed data by visit and change from baseline in ASQoL will be provided for each treatment.

[redacted]

PSQI

Summary statistics of observed data by visit and change from baseline in PSQI score will be provided for each treatment.

[redacted]

2.14 Interim analysis

The primary endpoint analysis was performed after all subjects completed Week 16 or discontinued earlier in order to support regulatory filing.

The investigators, site personnel and monitors continue to remain blinded to the treatment each subject received at randomization until all patients have completed the study (Week 60 Follow-up) and the final database lock has occurred. Additional analyses may be performed to support interactions with health authorities, as necessary.

3 Sample size calculation

An overall type I error (2-sided) of 5% will be used to control type I error. Secukinumab i.v. regimen will be tested versus placebo with respect to the primary endpoint (ASAS40 response at Week 16). A sample size of 250 axSpA patients (200 AS patients and 50 nr-axSpA patients) per group is deemed appropriate to achieve adequate power for the primary and secondary endpoints for this study.

An ASAS40 response rate of about 28.2% for the AS population in the placebo group at week 16 was reported in a published phase III study (CAIN457F2320). The ASAS40 response to secukinumab 150 mg is estimated to be 44.9% in the AS population based on the lower bound of the 90% confidence interval of risk difference from the Meta-analysis of 3 phase III studies (CAIN457F2310, CAIN457F2320 and CAIN457F2308). With 250 patients per treatment group (200 AS patients and 50 nr-axSpA patients), there would be approximately 96% power to detect a treatment difference of 16.7% in ASAS40 response rates between secukinumab and placebo in the evaluation of the primary efficacy hypothesis at Week 16 (Chi Square test, NQuery 7.0). The overall sample size will be 500 patients (400 AS patients and 100 nr-axSpA patients) for a randomization ratio of 1:1.

The estimated power with the chosen sample size for secondary efficacy endpoints based on the pooled results of the same phase III studies (CAIN457F2310, CAIN457F2320 and CAIN457F2308) are summarized in Table 3-1 for binary endpoints and Table 3-2 for continuous endpoints.

Table 3-1 Summary of power for binary secondary endpoints

Endpoint	Response Rate		Power
	Secukinumab(N = 200 AS + 50 nr-axSpA)	Placebo (N = 200 AS + 50 nr-axSpA)	
ASAS 5/6	44.2%	19.5%	99%
ASDAS-CRP major improvement	27.8%	6.4%	99%
ASAS20	59.0%	38.4%	99%
ASDAS-CRP inactive disease	13.6%	3.5%	97%
ASAS partial remission	15.2%	5.5%	92%

Table 3-2 Summary of power for continuous endpoints

Endpoint	Mean change from baseline		Common standard deviation	Power
	Secukinumab(N = 200 AS + 50 nr-axSpA)	Placebo (N = 200 AS + 50 nr-axSpA)		
hsCRP[1]	-0.79	0.09	0.867	99%
BASDAI	-2.61	-1.48	2.241	99%
BASFI	-2.14	-1.13	2.253	99%
SF-36 PCS	6.87	3.99	6.898	99%
ASQoL	-4.46	-2.56	4.714	99%
PSQI[2]	-1	0	3.9	81%

[1] hsCRP is in log scale
[2] Source: mean change from baseline is from Karatas et al 2018 (AS study, median was used) and common standard deviation is from Taylor-Gjevre et al 2011 (RA study, largest standard deviation was used).

4 Change to protocol specified analyses

No change from protocol specified analysis was made.

5 Appendix

5.1 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:

[redacted]

2. C-reactive protein (acute phase reactant)

ASAS40

The ASAS40 response is defined as a ≥ 40% improvement and an ≥ 2 units on a scale of 10 in at least three of the following 4 domains: back pain, patient global assessment of disease activity, physical function (BASFI) and inflammation (mean score of items 5 and 6 of the BASDAI) without any worsening in the remaining domain.

ASAS20

The ASAS20 response is defined as an improvement of ≥ 20% and ≥ 1 unit on a scale of 10 in at least three of the four main domains and no worsening of ≥ 20% and ≥ 1 unit on a scale of 10 in the remaining domain.

ASAS 5/6

The ASAS 5/6 improvement criteria is an improvement of ≥ 20% in at least five domains.

ASAS partial remission

The ASAS partial remission is defined as a value not above 2 units in each of the four main domains on a scale of 10.

[redacted] ASDAS-CRP, ASDAS Inactive disease, 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 C-reactive protein (CRP) in mg/liter ([redacted]) (Sieper 2009, Lukas 2009).

The ASDAS formulas are as follows:

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

[redacted]

CRP 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 are < 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 are 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 back pain and nocturnal back pain intensity (VAS)

The patient's assessment of 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

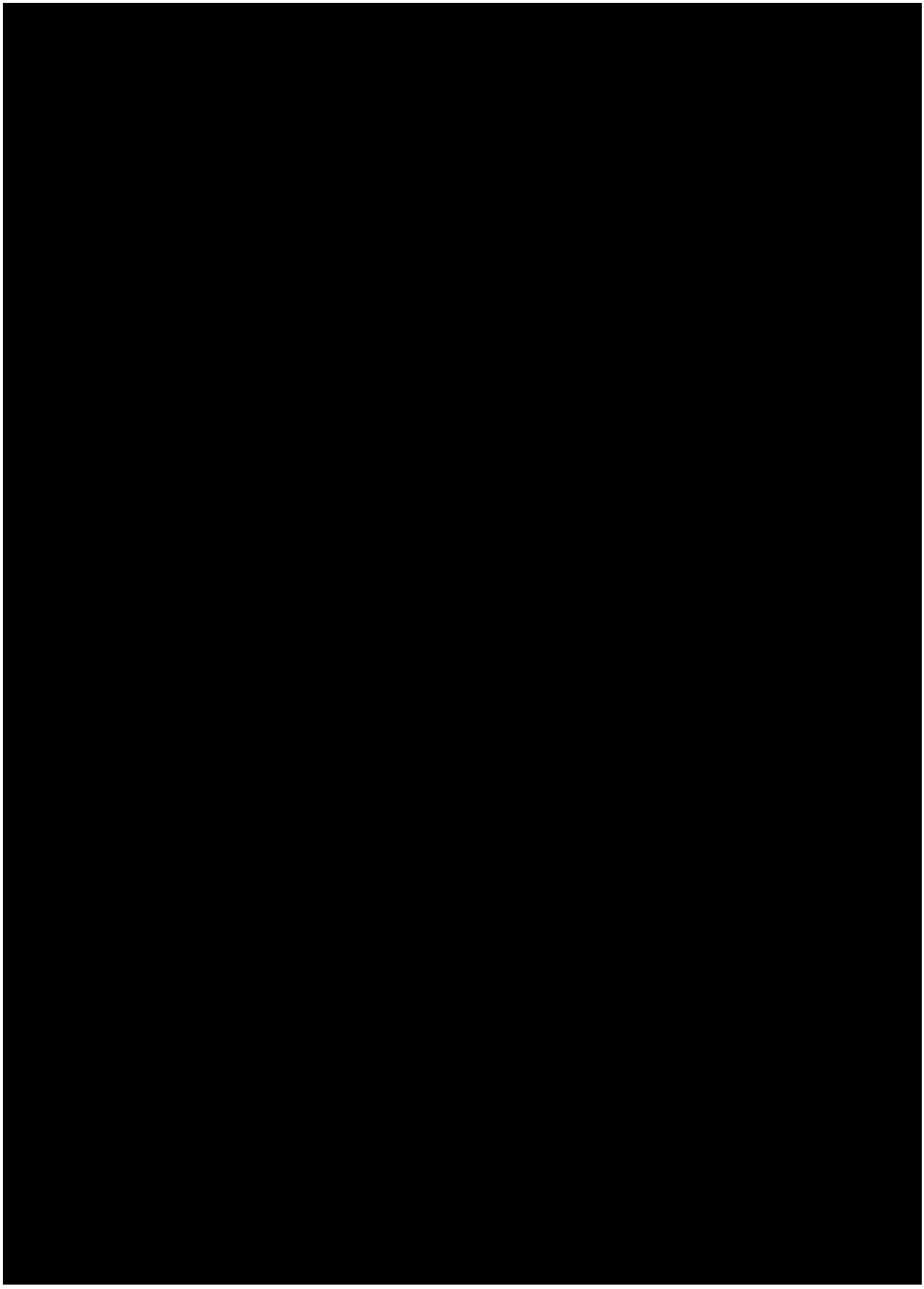
To give each symptom equal weighting, the mean (average) of the two scores relating to morning stiffness is taken and is then added to the sum of the first 4 questions. The resulting 0 to 50 score is divided by 5 to give a final 0 – 10 BASDAI score. Scores of 4 or greater suggest suboptimal control of 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. 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

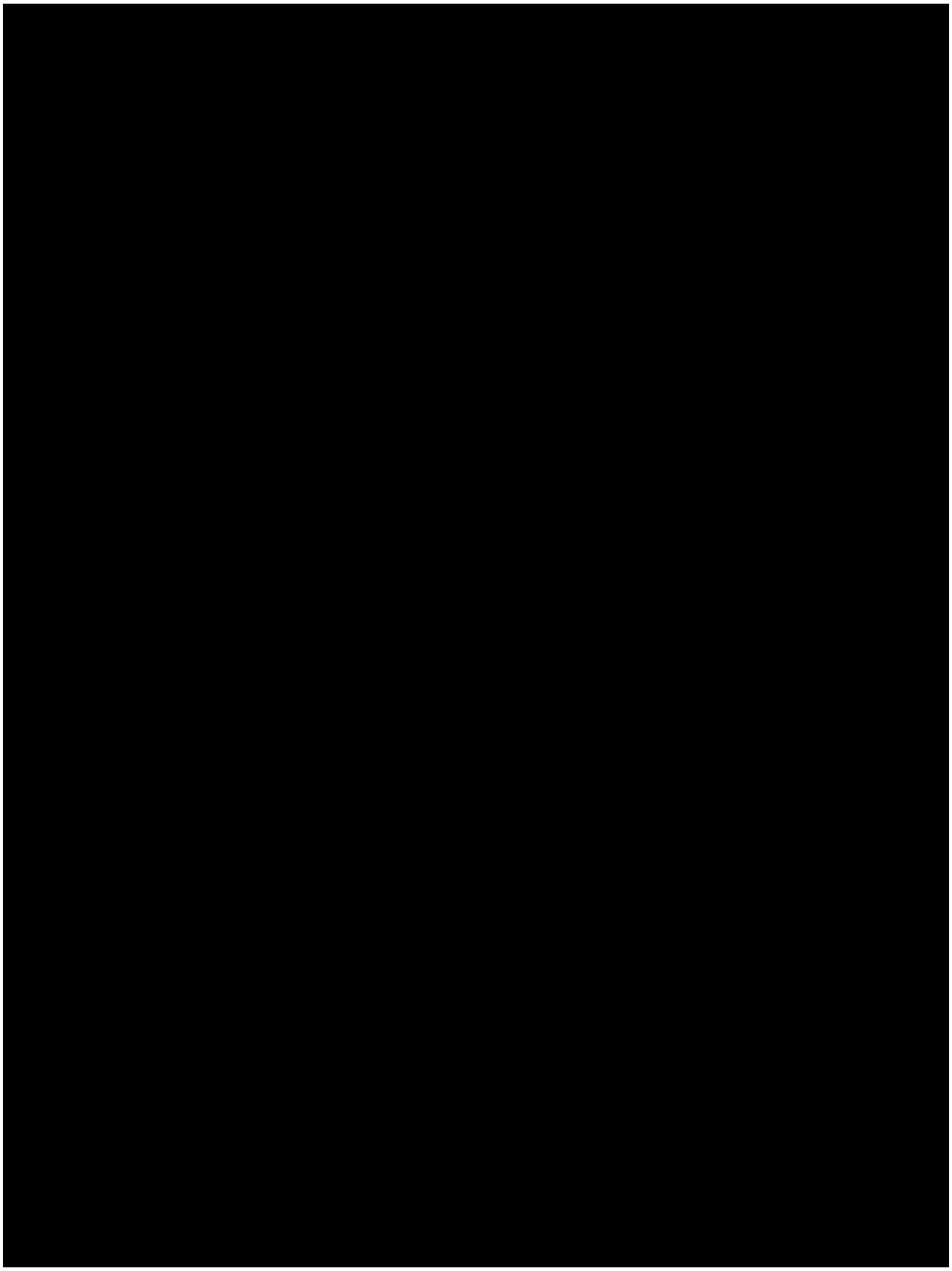
of 5. If two of Q1 to Q4 is missing and both Q5 and Q6 are not missing, then the sum should be divided by 3.

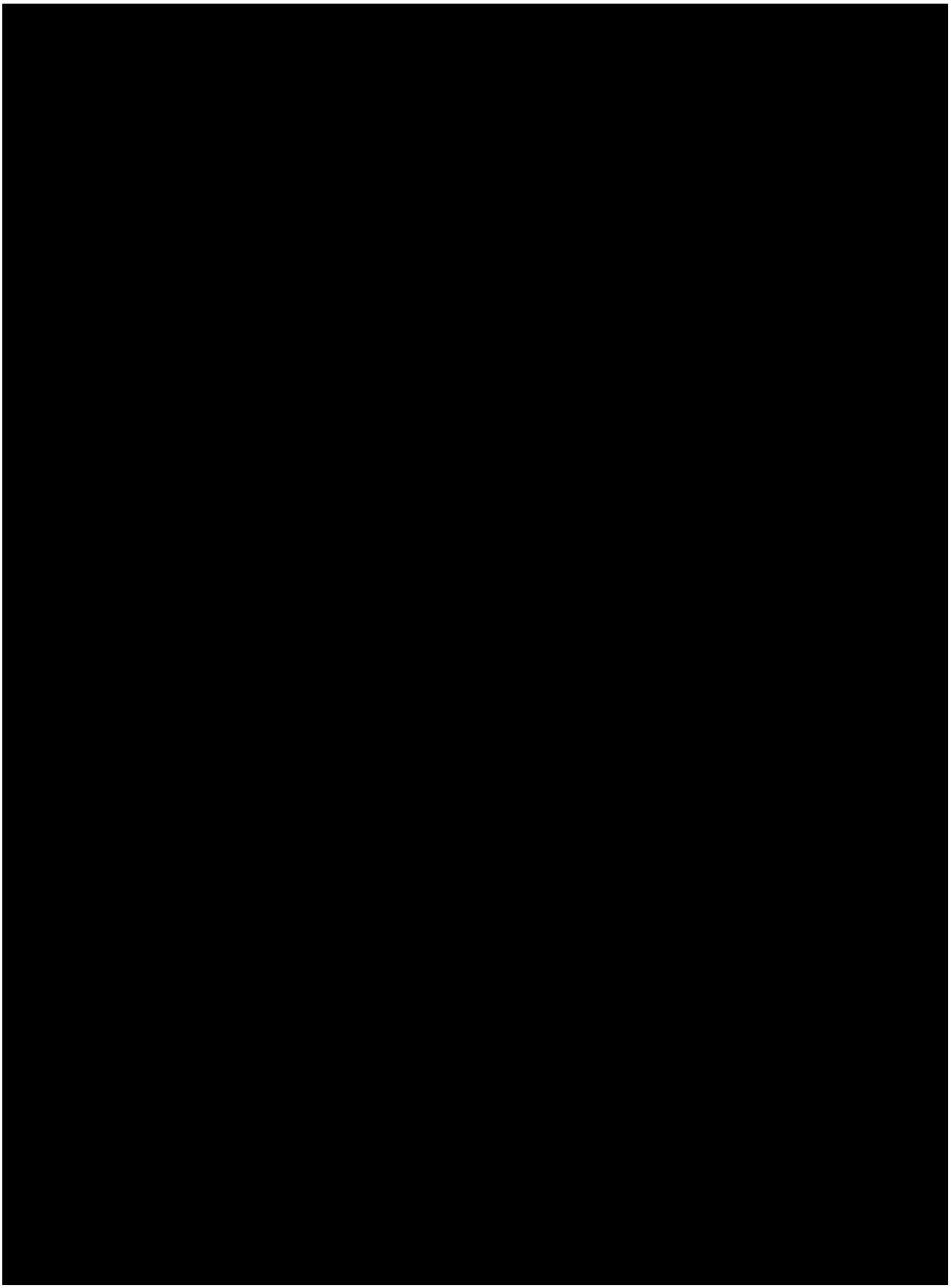
[redacted]

High Sensitivity C-reactive protein (hsCRP)

This assessment will be performed in order to identify the presence of inflammation, to determine its severity and to monitor the response to treatment.







[redacted]

5.2 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 indicates 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.

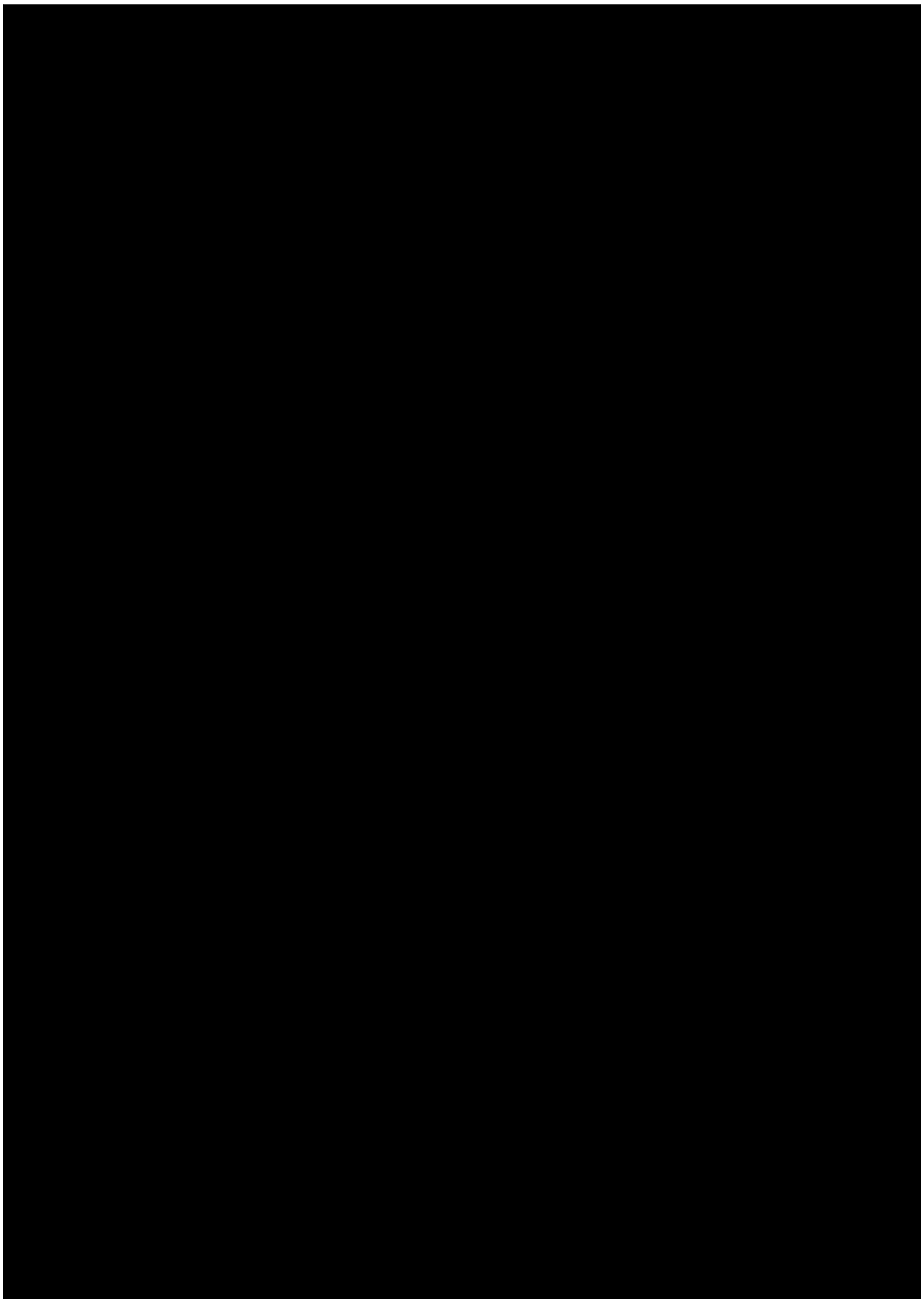
At least 15 answered questions are required to be able to calculate ASQoL using mean imputation, (sum of answered)/(number answered)*18 (Doward 2003).

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.

[redacted]



Pittsburgh Sleep Quality Index (PSQI)

The Pittsburgh Sleep Quality Index is a self-report questionnaire that assesses sleep quality over a 1-month time interval. Consisting of 19 items under 9 questions, the PSQI measures several different aspects of sleep, offering seven component scores and one composite score. The component scores consist of subjective sleep quality (1 item), sleep latency (2 items, i.e., how long it takes to fall asleep), sleep duration (1 item), habitual sleep efficiency (3 items, i.e., the percentage of time in bed that one is asleep), sleep disturbances (9 items), use of sleeping medication (1 item), and daytime dysfunction (2 items).

Of the 19 items, items 1-4 have free entry responses asking for usual bedtime and wake up time, number of minutes to fall asleep, and hours slept per night. Items 5-17 have 4-point Likert scale responses relating to frequency of specified sleep problems. Item 18 has a 4-pont Likert scale response relating to overall assessment of sleep quality (“very good”, “fairly good”, “fairly bad”, or “very bad”). Item 19 has a 4-point Likert response scale relating to respondent's overall

assessment of "enthusiasm to get things done" ("no problem at all", "only a very slight problem", "somewhat of a problem", or "a very big problem").

Each item is weighted on a 0-3 interval scale. In scoring the PSQI, seven component scores are derived, each scored 0 (no difficulty) to 3 (severe difficulty). The global PSQI score is then calculated by totaling the seven component scores, providing an overall score ranging from 0 to 21, where lower scores denote a healthier sleep quality and higher scores indicate worse sleep quality.

Questions 1 through 9 are not allowed to be missing. If these questions are missing then any scores calculated using missing questions are also missing.

5.3 Visit Windows

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 visit and 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.

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.

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

Table 5-4 Analysis visit windows

Analysis Visit	Target Day	Analysis Visit Window	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	Group 11	Group 12	Group 13
Baseline	1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1
Week 4	29	2-43	2-43	2-43			2-43	2-43	2-43	2-43	2-43	2-43	2-43		
Week 8	57	44-71	44-71	44-71	2-85		44-71	44-71	44-71	44-71	44-71	44-71	44-85		
Week 12	85	72-99	72-99	72-99			72-99	72-99	72-99	72-99	72-99	72-99			
Week 16	113	100-127	100-127	100-127	86-141	2-141	100-127	100-127	100-141	100-141	100-127	100-127	86-141	2-239	2-239
Week 20	141	128-155	128-155	128-155			128-239	128-155			128-155	128-155			
Week 24	169	156-173	156-183	156-183	142-267	142-267		156-197	142-197	142-197	156-183	156-183	142-253		
Week 25	176	174-187											**		
Week 28	197	188-211	184-211	184-211							184-211	184-211			
Week 32	225	212-239	212-239	212-239				198-253	198-253	198-253	212-239	212-239			
Week 36	253	240-267	240-267	240-267							240-267	240-267			
Week 40	281	268-295	268-295	268-295				254-323	254-323	254-323	268-295	268-295			
Week 44	309	296-323	296-323	296-323							296-323	296-323			
Week 48	337	324-351	324-351	324-351			240-351				324-351	324-379	***		
Week 52	365	352-393	324-393	352-435	268-435	268-435	352-435	324-393	324-393	324-393	352-393		352-393	240-393	240-435
Week 60	421	394-435	394-435					394-435	394-435	394-435	394-435	380-435	394-435	394-435	

Group 1: Patient's global assessment of disease activity (VAS), Patient's assessment of back pain intensity (VAS), BASFI, BASDAI, Vital signs [redacted]
Group 3: SF-36, ASQoL, [redacted], [redacted] Lipids
Group 4: [redacted], [redacted], Cardiovascular panel
Group 5: [redacted], PSQI, [redacted]
[redacted]
Group 7: hsCRP

Analysis Visit	Target Day	Analysis Visit Window	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	Group 10	Group 11	Group 12	Group 13
Group 8: hematology, blood chemistry, urinalysis Group 9: Urine pregnancy test Group 10: Weight Group 11: PK [redacted] * The first day of administration of randomized study treatment (first dose) is defined as Day 1. ** For Week 25 visit window, if the patient visited at week 24, then any visit at day range between 7-14 days after week 24 visit will be counted as Week 25; Otherwise, if the patient missed the visit of week 24, please use the range of 254-267 to count the visit into week 25. *** For Week 48 visit window, if the patient visited week 25 as defined above, then any day after week 25 visit up to day 351 would be counted as Week 48.															

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 5-5 Rules for flagging variables

Timing of measurement	Type of data	Rule
Baseline	All data	The last measurement made prior to administration of the first dose of study treatment – note this may include measurements taken on the day of randomization (e.g., lab). Baseline assessments scheduled for and captured on Day 1 will be considered baseline measurements regardless of the time of assessment. If a patient did not receive any dose of study treatment then the randomization date will be used. Only the date part will be considered if there is only one assessment on Day 1 but if there are multiple assessments on Day 1, then the following rules will apply: 1. If time of assessment exist, • select the last available measurement prior to the reference start date/time considering time • if no measurement prior to the reference start date/time then considering time select the earliest measurement post reference start date/time 2. If time of assessment does not exist the measurement from the lowest CRF visit number will be used.
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 from placebo to AIN within the window the following rules apply: 1. If the analysis visit window is ≤ week 16, then • 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 then the closest measurement after the switch to target will be used 2. If the analysis visit window is > week 16, then • If available, the closest measurement to the target date which is after the switch date will be used • If there are no data after the switch then the closest to target before or on the switch date will be use 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 use • 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, 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 from placebo to AIN within the window the following rules apply: 1. If the analysis visit window is ≤ week 16, then

Timing of measurement	Type of data	Rule
		• 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 then the closest measurement after the switch to target will be used 2. If the analysis visit window is > week 16, then • If available, the closest measurement to the target date which is after the switch date will be used • If there are no data after the switch then the closest to target before or on the switch date will be use 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 • If CRF visit number is the same the average value 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

5.4 Statistical methodology and assumptions

5.4.1 Analysis of continuous data

5.4.1.1 Summary statistics for continuous data

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

5.4.2 Analysis of binary and categorical data

5.4.2.1 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

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

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'.

5.4.3 Crude incidence and related risk estimates

5.4.3.1 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), 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

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

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.

5.4.4 Exposure adjusted incidence rate and related risk estimates

5.4.4.1 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 5-6 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

6 Reference

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.

[redacted]

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.

Karatas G, Bal A, Yuceege M, et al (2018) Evaluation of sleep quality in patients with ankylosing spondylitis and efficacy of anti-TNF-α therapy on sleep problems: A polisomnographic study. Int J Rheum Dis p. 1263-1269.

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 Anwer (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.

Taylor-Gjevre RM, Gjevre JA, Nair BV, et al (2011) Improved Sleep Efficiency after Anti-Tumor Necrosis Factor α Therapy in Rheumatoid Arthritis Patients. Ther Adv Musculoskelet Dis p. 227-33.

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

[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