STATISTICAL ANALYSIS PLAN

Study: AS0007

Product: Certolizumab pegol

MULTICENTER, OPEN-LABEL STUDY TO ASSESS THE EFFECTS OF CERTOLIZUMAB PEGOL ON THE REDUCTION OF ANTERIOR UVEITIS FLARES IN AXIAL SPONDYLOARTHRITIS SUBJECTS WITH A HISTORY OF ANTERIOR UVEITIS (C-VIEW)

SAP/Amendment Number	Date
Final SAP	26 May 2017
Amendment 1	09 May 2018
Amendment 2	14 Nov 2018
Amendment 3	21 Feb 2020

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Confidentiality Statement

Confidential

This document is the property of UCB and may not – in full or in part – be passed on, reproduced, published, or otherwise used without the express permission of UCB.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS 5
1 INTRODUCTION 9
2 PROTOCOL SUMMARY 9
2.1 Study objectives 10
2.1.1 Primary objective 10
2.1.2 Secondary objectives 10
2.1.3 Other objectives 10
2.1.4 Safety objective 10
2.2 Study variables 10
2.2.1 Efficacy variables 10
2.2.1.1 Primary efficacy variable 10
2.2.1.2 Secondary efficacy variables 10
2.2.1.3 Other efficacy variables 11
2.2.2 Safety variables 12
2.2.2.1 Secondary safety variable 12
2.2.2.2 Other safety variables 12
2.3 Study design and conduct 12
2.4 Determination of sample size 15
3 DATA ANALYSIS CONSIDERATIONS 16
3.1 General presentation of summaries and analyses 16
3.2 General study level definitions 17
3.2.1 Analysis time points 17
3.2.2 Relative day 17
3.3 Definition of Baseline values 18
3.4 Protocol deviations 18
3.5 Analysis sets 18
3.5.1 Enrolled Set 19
3.5.2 Safety Set 19
3.5.3 Full Analysis Set 19
3.5.4 Per Protocol Set 19
3.6 Treatment assignment and treatment groups 19
3.7 Center pooling strategy 19
3.8 Coding dictionaries 19
3.9 Changes to protocol-defined analyses 19
4 STATISTICAL/ANALYTICAL ISSUES 20
4.1 Adjustments for covariates 20
4.2 Handling of dropouts or missing data 20

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

4.3 Interim analyses and data monitoring 24
4.4 Multicenter studies 25
4.5 Multiple comparisons/multiplicity 25
4.6 Use of an efficacy subset of subjects 25
4.7 Active-control studies intended to show equivalence 25
4.8 Examination of subgroups 25
5 STUDY POPULATION CHARACTERISTICS 25
5.1 Subject disposition 25
5.2 Protocol deviations 26
6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS 26
6.1 Demographics 26
6.2 Other Baseline characteristics 26
6.3 Medical history and concomitant diseases 27
6.4 Prior and concomitant medications 28
7 MEASUREMENTS OF TREATMENT COMPLIANCE 28
8 EFFICACY ANALYSES 29
8.1 Statistical analysis of the primary efficacy variable(s) 30
8.1.1 Derivations of primary efficacy variable 30
8.1.2 Supportive and sensitivity analyses of the primary efficacy variable 31
8.2 Statistical analysis of the secondary efficacy variable(s) 32
8.2.1 Number of AU flares per 100 patient-years 32
8.2.2 Ankylosing spondylitis disease activity score (ASDAS) 32
8.2.3 Bath ankylosing spondylitis disease activity index (BASDAI) 33
8.2.4 ASAS20, ASAS40, and ASAS PR 34
8.2.5 Physician's global assessment of disease activity (PhGADA) 34
8.2.6 Swollen and tender joint counts 34
8.2.7 Patient's global assessment of disease activity (PtGADA) 35
8.2.8 Total and nocturnal spinal pain 35
8.2.9 Bath ankylosing spondylitis functional index (BASFI) 35
8.3 Analysis of other efficacy variables 36
8.3.1 Flare assessments 36
8.3.2 Extra-articular assessments 36
8.3.3 Ankylosing spondylitis quality of life (ASQoL) 36
8.3.4 ASAS Health Index (ASAS HI) Questionnaire 37
8.3.5 Short-Form 36-Item Health Survey (SF-36) 37
9 PHARMACOKINETICS AND PHARMACODYNAMICS 37
9.1 Pharmacokinetics 37
9.2 Pharmacodynamics 37

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

10 SAFETY ANALYSES 37
10.1 Extent of exposure 37
10.2 Adverse events 38
10.3 Clinical laboratory evaluations 40
10.4 Vital signs, physical findings, and other observations related to safety 42
10.4.1 Vital signs 42
10.4.2 Pregnancy testing 42
10.4.3 Physical assessments 42
10.4.4 MRI assessments 42
10.4.5 Tuberculosis assessments 42
10.4.6 Chest x-ray 43
11 REFERENCES 44
12 APPENDICES 45
12.1 APPENDIX 1: Compliance Ratio Calculation 45
13 AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN (SAP) (IF APPLICABLE) 47
13.1 AMENDMENT 1 47
13.2 AMENDMENT 2 54
13.3 AMENDMENT 3 63
STATISTICAL ANALYSIS PLAN SIGNATURE PAGE 92

LIST OF TABLES

Table 2–1: Schedule of assessments 13

LIST OF FIGURES

Figure 2–1: Schematic diagram 15

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

LIST OF ABBREVIATIONS

AE	adverse event
ALP	alkaline phosphatase
AS	ankylosing spondylitis
ASAS	Assessment of SpondyloArthritis international Society
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASQoL	Ankylosing Spondylitis Quality of Life
ATC	Anatomic Therapeutic Chemical
AU	anterior uveitis
axSpA	axial spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BMI	body mass index
BSA	body surface area
CII	clinically important improvement
CR	compliance ratio
CRD	compliance ratio based on duration
CRP	C-reactive protein
CZP	certolizumab pegol
DMARD	disease-modifying antirheumatic drug
EAER	exposure adjusted event rate
EAIR	exposure adjusted incidence rate
ES	Enrolled Set

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

FAS	Full Analysis Set
FDA	Food and Drug Administration
FU	Follow-Up
HI	Health Index
HLA-B27	human leukocyte antigen B27
HLT	high level term
HRQoL	health-related quality of life
IBD	inflammatory bowel disease
ICF	informed consent form
ICH	International Council for Harmonisation
ID	inactive disease
IGRA	interferon-gamma release assay
IPD	important protocol deviations
IWRS	interactive web response system
LLN	lower limit normal
LTBI	latent tuberculosis infection
MCID	minimal clinically important difference
MedDRA	Medical Dictionary for Regulatory Activities
MH	markedly abnormal high
MI	major improvement
ML	markedly abnormal low
mNY	modified New York (criteria)
MRI	magnetic resonance imaging
nr-axSpA	nonradiographic axSpA

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

NRS	Numerical Rating Scale
NSAID	nonsteroidal anti-inflammatory drug
PhGADA	Physician's Global Assessment of Disease Activity
PPS	Per Protocol Set
PR	partial remission
PT	preferred term
PtGADA	Patient's Global Assessment of Disease Activity
Q2W	every 2 weeks (every other week)
SAE	serious adverse event
SAP	Statistical Analysis Plan
sc	subcutaneous(ly)
SD	standard deviation
SF-36	Short-Form 36-Item Health Survey
SIJ	sacroiliac joint
SJC	swollen joint count
SpA	spondyloarthritis
SOC	System Organ Class
SS	Safety Set
TEAE	treatment emergent adverse event
TB	tuberculosis
TJC	tender joint count
TNF	tumor necrosis factor
ULN	upper limit normal
WD	withdrawal

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

WHO-DD	World Health Organization Drug dictionary

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

1 INTRODUCTION

This Statistical Analysis Plan (SAP) describes the analysis of study AS0007. It includes the final analysis when all data including follow-up period are available, as well as the formal interim analysis defined in the protocol, which will take place when all Week 48 data are available.

The SAP is based on Amendment 3 of the Clinical Study Protocol, dated 22 Jan 2020.

The content of this SAP is in compliance with the International Conference on Harmonization (ICH) / Food and Drug Administration (FDA) E9 Guidance documents.

2 PROTOCOL SUMMARY

AS0007 is a multicenter, open-label, Phase 4 study to evaluate the effects of certolizumab pegol (CZP) on the incidence of anterior uveitis (AU) flares per year in subjects with both active axial spondyloarthritis (axSpA) and a history of AU by comparing the historical AU flare rate that occurred prior to CZP treatment with the AU flare rate occurring while under CZP treatment.

Approximately 130 subjects will be screened in Europe in order to enroll 86 subjects into the study.

The study duration from the start of treatment will be 96 weeks from Baseline onwards, and a follow-up visit will be performed at Week 104; 10 weeks after the last dose administration at Week 94.

The study includes 3 periods as follows:

- **Period 1** (Screening Period): 1 to 5 weeks before Baseline

Subjects will be informed about the study and sign the informed consent form (ICF). Eligibility will be evaluated and assessments will be performed. The Screening Period should be kept as short as possible but can be extended to 5 weeks if certain medications need to be washed out or to allow to obtain information from the subject's ophthalmologist. For subjects who start a prophylactic treatment for latent tuberculosis infection (LTBI) the Screening Period can be extended up to 12 weeks.

- **Period 2** (Treatment Period): Week 0 to Week 96

Eligible subjects will receive a dose of CZP 400mg subcutaneously (sc) at Baseline, Week 2, and Week 4 followed by CZP 200mg sc every 2 weeks (Q2W) starting at Week 6 until Week 94.

All subjects will be trained at the beginning of the study on self-administration before the subjects start self-administration with the study drug (relative or caregiver may also perform the injections). The injection schedule will provide the sequence of self-administration and site visits including injection at the site.

- **Period 3** (Follow-Up [FU] Period): 10 weeks from the final dose of study medication received

All subjects will have a FU Visit at Week 104 or earlier in case of an early withdrawal (WD), 10 weeks after the final administration of CZP administration received within the study.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

A Week 48 analysis will be performed after all subjects have either completed the Week 48 Visit or have prematurely withdrawn prior to the Week 48 Visit. The final analysis will be conducted at the end of the study after all study data is locked.

2.1 Study objectives

The following key definition is provided as a reference for understanding the study objectives and is used throughout the document.

A flare is defined as being a new episode of AU that, based on the judgment of an ophthalmologist, requires specific treatment. A flare is considered a new episode if a gap of at least 3 months occurs between 2 flares.

2.1.1 Primary objective

The primary objective of the study will be to demonstrate the effect of CZP treatment on the reduction of AU flares in subjects with active axSpA and a documented history of AU.

2.1.2 Secondary objectives

The secondary objectives of the study will be to assess the effect of CZP treatment on:

- The reduction of AU flares in subjects with active axSpA having at least 1 documented history of AU within 12 months prior to Baseline
- AxSpA disease activity

2.1.3 Other objectives

The other objectives of the study will be to assess the effect of CZP treatment on:

- Physical function
- Signs and symptoms of axSpA
  - Morning stiffness
  - Fatigue
  - Extra-articular manifestations of axSpA
- Duration and severity of AU flares

2.1.4 Safety objective

The safety objectives of the study will be to evaluate the safety and tolerability of CZP therapy.

2.2 Study variables

2.2.1 Efficacy variables

2.2.1.1 Primary efficacy variable

The primary efficacy variable will be the count of distinct episodes of AU flares during the Treatment Period.

2.2.1.2 Secondary efficacy variables

The following secondary efficacy variables will be assessed at Week 48 and Week 96:

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Number of AU flares per 100 patient-years in subjects with active axSpA and a history of AU
- Number of AU flares per 100 patient-years in subjects with active axSpA and at least 1 AU episode within 12 months prior Baseline
- Change from Baseline in Ankylosing Spondylitis Disease Activity Score (ASDAS)
- Change from Baseline in Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)
- Assessment of ASAS 20, 40, 5/6, and partial remission (PR) response rates
- Change from Baseline in tender and swollen joint count (44 joint count); Physician's Global Assessment of Disease Activity (PhGADA).
- Change from Baseline in the respective components of the ASAS response criteria, assessed by:
  - Patient's Global Assessment of Disease Activity (PtGADA)
  - Pain assessment (total spinal pain Numerical Rating Scale [NRS])
  - Function (represented by Bath Ankylosing Spondylitis Functional Index [BASFI])
  - Inflammation (the mean of the BASDAI questions 5 and 6 concerning [redacted]

2.2.1.3 Other efficacy variables

The following other efficacy variables will be assessed at the timepoints indicated in Table 2–1 :

- Duration of AU flares
- Severity of AU flares
- Change from Baseline in ASDAS
- Change from Baseline in BASDAI
- ASAS 20, 40, 5/6, and PR response rates
- Change from Baseline in tender and swollen joint count (44 joint count); Physician's Global Assessment of Disease Activity (PhGADA)
- Change from Baseline in the respective components of the ASAS criteria, assessed by:
  - PtGADA
  - Pain assessment (total spinal pain NRS)
  - Function (represented by BASFI)
  - Inflammation (the mean of the BASDAI questions 5 and 6 concerning [redacted])
- Change from Baseline in ASDAS disease activity (clinical improvement [CI], major improvement [MI], inactive disease [ID], clinically important improvement [CII]), and BASFI) (including Weeks 48 and 96)

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Change from Baseline in [redacted] (NRS) (from BASDAI) (including Weeks 48 and 96)
- Change from Baseline in Ankylosing Spondylitis Quality of Life (ASQoL) (including Weeks 48 and 96)
- Change from Baseline in ASAS Health Index (HI) questionnaire (including Weeks 48 and 96)
- Change from Baseline in Short-Form 36-Item Health Survey (SF-36) (including Weeks 48 and 96)
- Number of inflammatory bowel disease (IBD) exacerbations
- Number of psoriasis exacerbations (in subjects with concurrent psoriasis)

2.2.2 Safety variables

2.2.2.1 Secondary safety variable

The secondary safety variable to be assessed is treatment-emergent adverse events (TEAEs).

2.2.2.2 Other safety variables

The following other safety variables to be assessed are:

- Blood pressure
- Physical examination
- Clinical laboratory values (hematology, biochemistry, and urinalysis).

2.3 Study design and conduct

AS0007 is a multicenter, open-label, Phase 4 study to evaluate the effects of CZP on the incidence of AU flares per year in subjects with both active axSpA and a history of AU by comparing the historical AU flare rate that occurred prior to CZP treatment with the AU flare rate occurring while under CZP treatment.

The study will be conducted in Europe. Approximately 130 subjects are planned to be screened in order to enroll 86 subjects into the study. The Schedule of Assessments is shown in Table 2–1. A study schematic diagram is presented Figure 2–1.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Table 2–1: Schedule of assessments

Study Period	Period 1 (Scree-ning)	Period 2 (Open-Label)												Period 3 FU
Week (W) Protocol Activity	**-5 to -1[a]**	**0 BL**	**2**	**4**	**12**	**24**	**32**	**36**	**48**	**60**	**72**	**84**	**96/ WD**	**104**
Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Written informed consent[b]	X													
Assessments for inclusion/ exclusion criteria	X	X												
Demographic data	X													
Medical history (including axSpA history, and tobacco use) and concomitant diseases	X													
Prior and concomitant medication[c]	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Uveitis history	X													
Uveitis flare evaluation		X	X	X	X	X	X	X	X	X	X	X	X	
Vital signs[d]	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Physical examination[e]	X	X	X	X	X	X		X	X	X	X	X	X	X
Extra-articular assessments[f]	X	X			X	X		X	X	X	X	X	X	
Hematology/ urinalysis/ biochemistry/	X[g]	X	X	X	X	X		X	X	X	X	X	X	
CRP	X	X	X	X	X	X	X	X	X	X	X	X	X	
Pregnancy testing[h]	X	X											X	X
Chest x-ray[i]	X													
TB test[j]	X								X				X	
TB questionnaire	X	X			X	X		X	X	X	X	X	X	
MRI (nr-axSpA subjects only)[k]	X													
X-ray (AS subjects only)[l]	X													
BASDAI	X	X	X	X	X	X	X	X	X	X	X	X	X	
BASFI		X	X	X	X	X		X	X	X	X	X	X	
ASQoL		X	X	X	X	X		X	X	X	X	X	X	
ASAS HI questionnaire		X	X	X	X	X		X	X	X	X	X	X	
SF-36		X	X	X	X	X		X	X	X	X	X	X	
PtGADA		X	X	X	X	X	X	X	X	X	X	X	X	

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Table 2–1: Schedule of assessments

Study Period	Period 1 (Scree-ning)	Period 2 (Open-Label)												Period 3 FU
Week (W) Protocol Activity	**-5 to -1[a]**	**0 BL**	**2**	**4**	**12**	**24**	**32**	**36**	**48**	**60**	**72**	**84**	**96/ WD**	**104**
Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Total and nocturnal spinal pain		X	X	X	X	X		X	X	X	X	X	X	
Swollen and tender joint counts		X	X	X	X	X		X	X	X	X	X	X	
PhGADA		X	X	X	X	X		X	X	X	X	X	X	
AEs		X	X	X	X	X	X	X	X	X	X	X	X	X
IWRS	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Study drug sc injections[m]		X	X	X	X	X	X	X	X	X	X	X		
Schedule appointment for next visit	X	X	X	X	X	X	X	X	X	X	X	X	X	

AE=adverse event; AS=ankylosing spondylitis; ASAS=Assessment of SpondyloArthritis international Society; ASQoL=Ankylosing Spondylitis Quality of Life; axSpA= axial spondyloarthritis; BASDAI=Bath Ankylosing Spondylitis Disease Activity Index; BASFI=Bath Ankylosing Spondylitis Functional Index; BL=Baseline; CRP=C-reactive protein; FU=Follow-up; HI=Health Index; HLA-B27= human leukocyte antigen B27; IBD=inflammatory bowel disease; IWRS=Interactive Web Response System; LTBI=latent tuberculosis infection; mNY=modified New York (classification criteria); MRI=magnetic resonance imaging; nr-axSpA=nonradiographic axSpA; PhGADA=Physician's Global Assessment of Disease Activity; PtGADA=Patient's Global Assessment of Disease Activity; Q2W=every 2 weeks (every other week); sc=subcutaneous(ly); SF-36=Short-Form 36-Item Health Survey; SIJ=sacroiliac joint; TB=tuberculosis; WD=Withdrawal

[a] For subjects with LTBI who initiate a prophylactic treatment and need to have 8 weeks of prophylactic treatment prior to randomization, the Screening Period can be extended to 12 weeks. For other subjects, the Screening Period should be kept as short as possible, but can be up to 5 weeks if certain medications need to be washed out or to allow to obtain information from the subjects' ophthalmologist.

[b] Informed consent: prior to any study activities, subjects will be asked to read and sign the informed consent form.

[c] Past medications will recorded at Screening. Concomitant medications will be recorded at all other visits.

[d] Pulse rate, systolic and diastolic blood pressures, and temperature are to be measured at Screening and Baseline, thereafter systolic and diastolic blood pressures are to be measured.

[e] Weight is to be measured at Screening, Baseline, Week 48, and at completion at Week 96/WD Visit. Height will be measured at the Baseline Visit only.

[f] Extra-articular assessments includes the number of IBD exacerbations and number of psoriasis exacerbations (in subjects with concurrent psoriasis)

[g] Testing to rule out hepatitis B core antibody, hepatitis B surface antigen, hepatitis B surface antibody, antibodies to hepatitis C, and antibodies to human immunodeficiency virus at Screening only. Testing for HLA-B27 at Screening, only if not performed before.

[h] Pregnancy testing for women of childbearing potential will be serum testing at the Screening Visit and urine dipstick testing at Baseline and Week 96/WD Visit and the FU Visit (10 weeks after the final dose of study drug).

[i] Chest x-ray used for screening must have been done within 3 months prior to the Screening Visit and should be repeated only if the TB test was confirmed positive or any further evidence is suggestive of potential TB infection (eg, exposure). If no chest x-ray is available within the 3 months prior to the Screening Visit, the x-ray can be done during the Screening Period.

[j] QuantiFERON TB GOLD test. The TB tests will be repeated at Week 48 and 96 (or at WD Visit if medically indicated) for subjects with previously negative TB test result.

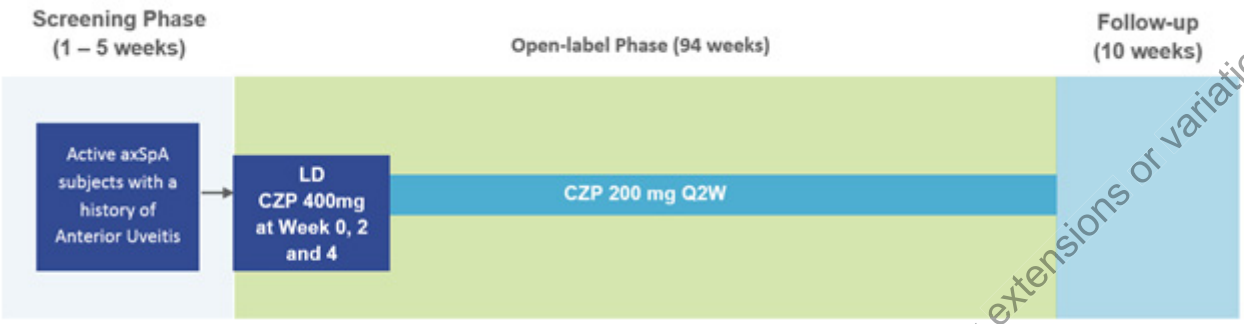
[k] For subjects with nr-axSpA, an MRI of the SIJs is to be performed during the Screening Period. An MRI performed ≤3 months prior to the Baseline Visit may be used as the Baseline. An MRI assessment is not needed for patients with AS.

[l] AS subjects must have evidence of sacroiliitis on x-ray taken prior to Baseline meeting the mNY classification criteria according to the Investigator. If not available, this x-ray is to be taken during the Screening Period. If an x-ray was performed prior to the Screening Visit and the mNY classification criteria were met, the x-ray is not to be repeated.

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
RELEASED COPY

[m] Onsite study drug administration will occur after all other visit assessments are completed and laboratory samples are drawn. Study drug injections will occur at home Q2W on Weeks 6, 8, and 10, 14 to 22, 26 to 30, 34, 38 to 46, 50 to 58, 62 to 70, 74 to 82, and 86 to 94. At all visits, an even number of syringes (2 per box) will be assigned (eg, at Week 4 syringes will be assigned for Weeks 6, 8, and 10; 1 syringe will be returned by the subject at Week 12).

Figure 2–1: Schematic diagram



axSpA=axial spondyloarthritis; CZP=certolizumab pegol; LD=loading dose; Q2W=every 2 weeks

The aim of the study is to evaluate the flare rate of AU before the start of CZP treatment compared to that of a period with CZP treatment. Since the occurrence of flare is both subject and disease specific, each subject should act as his/her own control. Because an untreated uveitis flare can cause severe eye complication, it was deemed necessary to treat all subjects with an active drug.

Therefore, 1 treatment arm with 1 set of subjects is deemed the right approach to allow for the comparison of the history of flare occurrence with the flare occurrence on a biological treatment in the same subjects. The dose of CZP used in this study is the same as per label to treat axSpA subjects, which has been shown to be [redacted] and effective for reducing the main axSpA symptoms.

2.4 Determination of sample size

Based on results from a previous study with CZP in axSpA subjects (AS001), the AU flare rate while treated with CZP is estimated to be 14.6 events per 100 subject-years of exposure. Assuming the rate while treated with CZP represents a reduction in the flare rate of approximately 50%, the AU flare rate for the prestudy period prior to initiation of CZP is estimated to be 29.2 events per 100 subject-years of exposure. A conservative estimate of the sample size for this study is obtained via derivation of the per treatment-group sample size for Poisson regression based on the approach for a 2-group setting initially described by Signorini (1991) and implemented within SAS (Hu, 2008). Assuming an average follow-up period of 1.5 years with 2-sided α-level=0.049 for the final analysis of the primary efficacy variable, a sample size of N=86 subjects per treatment group will have 80% statistical power for a 2 group comparison, corresponding to a total sample size of N=86 subjects for this study with 1 treatment group. Due to the expected increase in statistical efficiency from the fact that each subject serves as his/her own control in the primary efficacy analysis, the actual power for this study is assumed to be >80 percent.

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

3 DATA ANALYSIS CONSIDERATIONS

3.1 General presentation of summaries and analyses

The statistical analyses will be performed by Chiltern. Computations and generation of outputs will be performed using SAS® (Statistical Analysis System, SAS-Institute, Cary, NC, USA) Version 9.3 or above. All tables and listings will use Courier New font size 9.

For continuous data in general, summary statistics (n [number of available measurements], arithmetic mean, standard deviation (SD), median, minimum, and maximum) will be presented. For selected variables, the first (25th percent) quartile (Q1) and the third (75th percent) quartile (Q3) may also be presented (for patient reported outcomes).

Mean, SD, median and quartiles will be displayed to 1 more decimal place than collected in the source data. Derived variables in general will display the mean, SD and median to 1 more decimal place than the variables used in the derivation (i.e. a change from Baseline will be reported to the same precision as the Baseline data). However, a special rule can be defined as needed for appropriate reporting.

For descriptive statistics of continuous variables by visit, the change from Baseline and actual value at the given time point will be displayed. The change from Baseline is the post-Baseline value minus the Baseline value. If the Baseline or post-Baseline value is missing, then the change from Baseline is set to missing. Percent change from Baseline is the change from Baseline divided by Baseline and multiplied by 100. If the Baseline value is 0 and the post-Baseline value is also 0, then the percent change from Baseline is set to 0. If the Baseline value is 0 and the post-Baseline value is non-zero, then the percent change from Baseline is set to missing. If the Baseline value is missing, the percent change from Baseline is set to missing.

Frequency tables (frequency counts and percentages) will be presented for categorical data. If there are no missing values then the missing row can be removed. If there are missing values (including missing a single assessment or entire visit), then include the missing row with the frequency count and percentage. If there are no subjects in a specific CRF category, then that row will be retained and 0 presented in the table. In general, percentages will be calculated based on the utilized analysis set. However, in the case of subgroup analyses, the N of the subgroup will be used as denominator.

Unless stated otherwise, all inferential statistical tests will be 2-sided and conducted at the 0.05 alpha level. Adjustment for multiplicity is described in Section 4.5. P-values for secondary efficacy variables are nominal. P-values will be presented to 3 decimal places and any values below 0.001 should be displayed as “<0.001”.

All data in the database (SDTM and ADaM) will be presented in by-subject data listings, and sorted by subject number, parameter (where applicable), and visit (where applicable). For listings including subjects that were not treated, these subjects will be shown first in the listing, ordered by subject number. All listings will include repeated and unscheduled measurements. Such measurements will appear in chronological order together with the scheduled visits, i.e., a repeated measurement will appear directly after the visit and time relative to dosing, for which the repeat measurement was performed. In all the listings dates will be presented in the format ‘YYYY-MM-DD’ and times will be presented in 24h clock format as ‘hh:mm’ or ‘hh:mm:ss’ where appropriate.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

All tabulations will be sorted by parameter and visit. Only scheduled visits will be included in the tabulation. Categorical data will be summarized by visit, using the number and percentage of subjects in each category. Percentages will be based on the corresponding population size (i.e., the denominator of percentages should match the sample size in the column header), unless otherwise noted via footnote in the applicable summary table. Percentages will be presented to 1 decimal place. For data points with n=0 (i.e., no subjects in the applicable category), no value for percentage of subjects will be displayed.

3.2 General study level definitions

3.2.1 Analysis time points

The study includes 3 periods as follows:

- **Period 1** (Screening Period): 1 to 5 weeks before Baseline

The pre-treatment period (Screening period) of the study is the period prior to a subject's first dose of study medication intake. This period starts at the Screening visit (week -5 to -1 day, extended to week -12 for subjects who start a prophylactic treatment for LTBI) and ends at the Baseline visit (Week 0) up to the time of first study medication administration (exclusive). Unless specific time information is available to indicate that a Screening or Baseline visit assessment was performed after a subject's first study medication administration, most assessments will be attributed to the Screening period. Exceptions to this are AEs that start on Baseline visit, as they will be attributed to the Treatment Period.

- **Period 2** (Treatment Period): Week 0 to Week 96

The treatment period begins at the Baseline visit (Week 0) at the time of first study medication administration (inclusive) and ends at Week 96/WD visit. Premature withdrawal visit assessments will be assigned to the next scheduled visit following the last visit where assessments are available. All visit measurements, even violating the visit window, will be utilized for the respective visit as long as they are in the proper sequence. Subjects will be classified as completing the study if they complete the Week 96 Visit without early withdrawal of the study. This is regardless of whether they attend the Follow-up visit.

- **Period 3** (FU Period): 10 weeks from the final dose of study medication received

The Follow-up visit will take place 10 weeks after the last dose of study medication. That is either 8 weeks after the Week 96 visit for subjects completing the study, or up to 10 weeks after the WD visit for withdrawn subjects.

3.2.2 Relative day

The relative day will be included in different listings and will be calculated as follows:

- If the start (stop) date occurred on or after the first dose, but prior to the drug stop date, relative day is calculated as start (stop) date minus first dose date + 1.
- If the start (stop) date occurred after the last dose of drug, the relative day to the most recent dose is calculated as start (stop) date minus most recent dose date. The relative day in this situation should be preceded by a '+'.

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

- If the start (stop) date occurred before the first dose, the relative day is calculated as start (stop) date minus first dose date. The relative day in this situation should be preceded by a '-'.

For AEs, relative days for start and stop dates will be calculated as the number of days since the first injection of the medication. For non-treatment emergent AEs, relative day of onset will be negative if the event started and stopped before the first dose. Relative day will only be computed for fully completed dates and will be missing for partial dates.

3.3 Definition of Baseline values

Unless otherwise specified, the last valid measurement before study medication administration will be utilized as Baseline value. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables, assessments may be scheduled for Screening only and not for Baseline. In this case the screening value will be utilized as baseline value. If a baseline measurement is missing, and a screening value available, the screening value will be utilized as Baseline instead. If a scheduled Baseline assessment is taken on the same day as first administration of study medication, it will be used as Baseline, irrespective of the time of the assessment relative to the first administration of study medication.

3.4 Protocol deviations

Important protocol deviations (IPD) are deviations from the protocol which could potentially have a meaningful impact on study conduct, on the safety for an individual subject, or on the primary efficacy outcome for an individual subject. The criteria for identifying important protocol deviations will be defined within the specification of the IPD document. Important protocol deviations will be reviewed as part of the ongoing data cleaning process and all important deviations will be identified and documented prior to database lock.

Important protocol deviations will be identified and classified by the deviation types defined in the IPD document. A listing of all important protocol deviations identified at the Data Evaluation Meeting will be presented for all subjects in the SS, and will include deviation type and description. The number and percentage of subjects in the SS with important protocol deviations will be summarized overall.

The following deviation types, and further deviation types, may be included in the IPD:

- Inclusion criteria
- Exclusion criteria
- Withdrawal criteria
- Exposure
- Non-compliance
- Prohibited concomitant medications

3.5 Analysis sets

The main analysis sets defined for this study are the Enrolled Set (ES) and the Safety Set (SS). In addition, a Full Analysis Set (FAS) is defined for the unexpected case of missing Baseline values

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

for the primary efficacy variable. A Per Protocol Set (PPS) is added for additional analysis of the primary efficacy variable only.

3.5.1 Enrolled Set

The ES will consist of all subjects who have given informed consent.

3.5.2 Safety Set

The SS will consist of all subjects in the ES who have received at least 1 dose of study medication.

Safety variables, and also efficacy variables, will be summarized using the SS.

3.5.3 Full Analysis Set

The FAS will consist of all subjects in the SS with non-missing Baseline value for the primary efficacy variable, i.e. AU flare incidence data from pre-study period.

In case the FAS differs from the SS, the efficacy analysis will be conducted using the FAS instead of the SS.

3.5.4 Per Protocol Set

The PPS will consist of all subjects in the SS, or in the FAS if applicable, which do not have any important protocol deviation affecting the primary efficacy variable. Subjects with at least one important protocol deviation affecting the primary efficacy variable will be excluded from the PPS.

3.6 Treatment assignment and treatment groups

At Baseline, eligible subjects will be assigned to the study treatment CZP. Certolizumab pegol will be administered as 400mg at Baseline and Weeks 2 and 4 followed by 200mg Q2W (starting at Week 6 until Week 94). As there is a single treatment arm in this study, all subjects will be presented together in a single treatment group (eg, as “CZP”). For presentation purposes, the term CZP will be used.

3.7 Center pooling strategy

All results will be summarized across centers and will not be stratified by center. No pooling of centers is planned for this study.

3.8 Coding dictionaries

All AEs, concomitant diseases, and medical history will be coded for analysis according to the Medical Dictionary for Regulatory Activities (MedDRA)® coding dictionary, version 19.0. Prior and concomitant medications will be coded for analysis using the World Health Organization Drug dictionary (WHO-DD), version SEP/2015.

3.9 Changes to protocol-defined analyses

The analysis of Axial SpondyloArthritis international Society response ASAS5/6 will not be performed as described in the clinical study protocol, because one main component for spinal mobility, the Bath Ankylosing Spondylitis Metrology Index (BASMI), is not collected in the data.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

The ASAS 5/6 response is defined as at least 20% improvement in 5 of the following 6 domains:

- Patient's Global Assessment of Disease Activity
- Pain assessment (the total spinal pain NRS score)
- Function (represented by the Bath Ankylosing Spondylitis Functional Index (BASFI))
- Inflammation (the mean of the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) questions 5 and 6) [redacted])
- Spinal mobility (lateral lumbar flexion measured by BASMI)
- CRP as more objective measures

As the BASMI is not collected, and there is no alternative measure of spinal mobility available in the study data, the complete component for spinal mobility is missing. Therefore the ASAS 5/6 response criterion cannot be calculated, and the analysis of the secondary efficacy variable ASAS 5/6 has to be dropped.

The clinical study protocol does not mention a PPS. This analysis set has been added, in order to additionally perform the primary efficacy analysis on the subset of protocol-compliant subjects.

In the calculation of ASDAS (Section 9.1.2 of the protocol), the CRP component was written as "0.579 × (natural logarithm of the CRP [mg/L] + 1)", but should have been "0.579 × (natural logarithm of the (CRP [mg/L] + 1))".

For BASDAI (Section 9.1.3 of the protocol), the protocol stated "The BASDAI is a validated self-reported instrument which consists of six 10-unit horizontal NRSs to [redacted] over the last week". However the severity and duration are only measured for [redacted], the others are measured for severity only.

In Section 9.1.4 of the protocol, the pain assessment component of ASAS20/40/PR was defined as "the average of total and nocturnal spinal pain NRS scores", but this should be defined as total spinal pain NRS score only.

ASDAS Clinical improvement was detailed as an "other" efficacy variable, but no definition was given. As there is no well-accepted definition, this variable has not been included in the analysis.

In Section 12.7 of the protocol, it stated that the interim analysis would use α=0.0001 and so the final analysis would use α=0.049. This should have said that α=0.001 would be used for the interim analysis.

4 STATISTICAL/ANALYTICAL ISSUES

4.1 Adjustments for covariates

Statistical models may be adjusted for covariates. Any such adjustments will be described in the context of the analyses to be performed.

4.2 Handling of dropouts or missing data

Analysis is based on observed case only.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
EMA - CTED COPY

If the Baseline value is missing, the Screening value may be used as the Baseline. In general, if no measurement is available prior to receiving study medication, then the Baseline value is treated as missing, and will not be imputed.

Premature withdrawal visits should be assigned to what would have been the next scheduled visit for the particular assessment that is being summarized. If a premature withdrawal visit occurs on the same date as a scheduled visit, then the assessments will be summarized with the scheduled visit.

Safety variables

For analyses of AEs and concomitant medication usage, a complete date must be established in order to correctly identify the AE or medication as occurring during treatment or not. For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the subject data listings (i.e., no imputed values will be displayed in data listings).

- Missing start day, but month and year present: If the start of study medication occurred in the same month and year as the occurrence of the AE/concomitant medication, the start day of the event/concomitant medication will be assigned to the day of first intake of study medication. Otherwise the start day will be set to the 1st day of the month.
- Missing start day and month, but year present: If the start of study medication occurred in the same year as the occurrence of the AE/concomitant medication, the start day and month will be assigned to the date of first intake of study medication. Otherwise the start day and month will be set to January 1st.
- If start day, month, and year are missing, start date will be replaced by the date of first study drug administration or by the first day of January of the year of the end date, if the first administration date is after the end date.
- Missing end day, but month and year present: The end day will be set to the last day of the month.
- Missing end day and month, but year present: If the maximum of the subject's date of study termination or the date equivalent to 70 days after the subject's last intake of study medication is the same year as the occurrence of the AE/Concomitant medication, then the end day and month will be set to the maximum of the date of study termination or the date equivalent to 70 days after last intake of study medication. Otherwise the end day will be set to December 31st of the given year.

Efficacy variables

The start and end dates of flares will be imputed for the analyses of AU flares data, in order to allow the calculation of time intervals between 2 consecutive flares and the duration of flares. Different rules will be applied for pre-study and on-study period flares.

For the interim analysis, the pre-study period is starting 48 weeks prior to Baseline, i.e. 336 days prior to Baseline until Baseline. The on-study period for the interim analysis is starting from Baseline until week 48, i.e. until study day 336.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

For the final analysis, the pre-study period is starting 24 months prior to Baseline, i.e. 730 days prior to Baseline until Baseline. The on-study period for the final analysis is starting from Baseline until week 96, i.e. until the date of the Week 96 or discontinuation visit. In addition, sensitivity and supportive analyses use different definitions (as defined in Section 8.1.2).

Pre-study period:

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Month and year are the same as month and year of the end date of previous flare in the same eye, and the end date is clearly greater than end date of previous flare, then impute with the end date of previous flare.
  - In the case that month and year of the start date of pre-study period is the same month and year of the end date of previous flare in the same eye, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Year is the same as year of the end date of previous flare in the same eye, then impute with the end date of previous flare.
  - In the case that year of the start date of pre-study period is the same year of the end date of previous flare in the same eye, then use the latest date of both for imputation.
- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred within the pre-study period. Then the maximum (latest) of start date of pre-study period and end date of previous flare will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Month and year are the same as month and year of the start date of subsequent pre-study flare in the same eye before Baseline, then impute with the start date of subsequent pre-study flare.
  - In the case that month and year of the end date of pre-study period is the same month and year of the start date of subsequent pre-study flare in the same eye, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- Year is the same as year of the end date of pre-study period, then impute with the end date of pre-study period.
- Year is the same as year of the start date of subsequent pre-study flare in the same eye, then impute with the start date of subsequent pre-study flare.
- In the case that year of the end date of pre-study period is the same year of the start date of subsequent pre-study flare in the same eye, then use the earliest date of both for imputation.

- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred within the pre-study period (of 2 years prior to baseline). Then the minimum (earliest) of end date of pre-study period (i.e. the day prior to treatment start date), the day prior to start date of subsequent pre-study flare in the same eye, or the start date plus the median duration of flare in that subject will be considered. The median duration of flare is determined from all flares recorded for that subject that have complete start and end dates (regardless of eye), including any flares prior to or during the pre-study period and on-study period.

On-study period (Baseline until end of treatment):

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of study medication, then impute with the start date of study medication.
  - Month and year are the same as month and year of the end date of previous on-study flare in the same eye, and the end date is clearly greater than end date of previous on-study flare in the same eye, then impute with the end date of previous on-study flare in the same eye.
  - In the case that month and year of the start date of study medication is the same month and year of the end date of previous on-study flare in the same eye, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of study medication, then impute with the start date of study medication.
  - Year is the same as year of the end date of previous on-study flare in the same eye, then impute with the end date of previous on-study flare in the same eye.
  - In the case that year of the start date of study medication is the same year of the end date of previous on-study flare in the same eye, then use the latest date of both for imputation.
- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred after treatment start during the on-study period. Then the maximum (latest) of start date of study medication and end date of previous on-study flare in the same eye will be considered.

This document cannot be used to support any marketing authorization application and/or extensions or variations thereof.
REDACTED COPY

- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of on-study period, then impute with the end date of on-study period.
  - Month and year are the same as month and year of the start date of subsequent on-study flare in the same eye, then impute with the start date of subsequent on-study flare in the same eye.
  - In the case that month and year of the end date of on-study period is the same month and year of the start date of subsequent on-study flare in the same eye, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of on-study period, then impute with the end date of on-study period.
  - Year is the same as year of the start date of subsequent on-study flare in the same eye, then impute with the start date of subsequent on-study flare in the same eye.
  - In the case that year of the end date of on-study period is the same year of the start date of subsequent on-study flare in the same eye, then use the earliest date of both for imputation.
- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred after treatment start during the on-study period. Then the start date of subsequent on-study flare in the same eye will be considered. Otherwise it will be assumed to be ongoing.

4.3 Interim analyses and data monitoring

Regular monitoring of safety data collected during clinical studies will be performed as described in the Safety Signal Detection in the Ongoing Clinical Trials Charter for CZP.

A specific data monitoring, steering, or evaluation committee is not planned for this study.

An interim analysis is planned to be conducted after the last subject has completed the Week 48 assessment. That is when each treated subject has completed the Week 48 Visit, or withdrawn prematurely. All data collected until the database cut-off date will be included. With this interim analysis, first study results will be available already about one year before the final analysis of the study. There is no intention to stop the study early for efficacy or for futility. The interim summaries and listings will be based on the same shells as for the final analysis, unless otherwise stated. The interim listings will include all data collected up to Week 48. The summary tables presenting results by visit will only consider results up to Week 48 Visit for the interim analysis. Summary tables of results not provided by visit, e.g. adverse events, will not include events or medication that started after the Week 48 Visit. Data collected after Week 48 will not be presented in the interim analysis.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

The length of pre-study period and on-study period for collection of AU flares data will be 48 weeks for the interim analysis, and will be 96 weeks for the final analysis. The analysis based on period length of 48 weeks will be repeated for the final analysis.

No separate SAP will be created for the interim analyses, as all analyses will be covered within this SAP, and within the corresponding display specifications. In general, the interim analysis will follow the same specifications as the final analysis, unless stated differently in the corresponding sections. The interim analysis will include a subset of tables and listings that will be marked in the display specifications.

An additional interim data-cut might be needed to support a regulatory submission in nr-axSpA or axSpA.

4.4 Multicenter studies

The data from all centers will be pooled for the purposes of the final analyses. There will be no formal statistical evaluation of the effect of center on the results obtained.

4.5 Multiple comparisons/multiplicity

Although there is no intention to stop the study early due to efficacy or futility on the basis of the interim results, α=0.001 will be spent in conjunction with this formal interim review of the data (Haybittle, 1971), and the final analysis of the primary efficacy variable will be conducted at the reduced 2-sided α-level of 0.049.

4.6 Use of an efficacy subset of subjects

Not applicable.

4.7 Active-control studies intended to show equivalence

Not applicable.

4.8 Examination of subgroups

The primary efficacy variable will be summarized using the following subgroups:

- Gender (Male, Female)
- Age (<median, >=median)
- axSpA subpopulation as determined by the investigator (Ankylosing Spondylitis, non-radiographic axial spondyloarthritis)

A Poisson regression model as described for the primary analysis will be used including terms for the subgroup and the subgroup by treatment period interaction. The results will be presented, but no p-values will be given.

5 STUDY POPULATION CHARACTERISTICS

5.1 Subject disposition

Subject disposition will be listed for all subjects screened and will include the following information: subject status (screen failure, completed or discontinued), date of informed consent, date of first and last dose of study medication (including relative day for last dose), date of premature study termination (if applicable) and primary reason for study termination (if

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

applicable). For screen failures the date and reason for screen failure will be listed. Subject disposition will be summarized for all subjects screened including the reasons for screen failure. Subject disposition will also be summarized for the Enrolled Set including the reasons for discontinuation of study and/or study medication. Study eligibility criteria will be listed and a separate listing of subjects who did not meet the eligibility criteria will be presented. A listing of study visit dates will be presented by subject including the relative study day with respect to treatment start for each visit.

5.2 Protocol deviations

Protocol deviations will be identified and documented as described in Section 3.4. The assignment of subjects to each of the analysis sets will be listed for all subjects screened. In addition, a listing of subjects excluded from the analysis sets will be presented including the reasons for exclusion; this listing will be based on the Enrolled Set.

6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

6.1 Demographics

Tables with descriptive statistics and listings will be given for the demographic variables age (at time of informed consent), gender, weight and height at Baseline, race, ethnicity, body mass index (BMI), and body surface area (BSA) as calculated below. Demographic characteristics will be summarized for the SS.

Only the year of birth has been collected, so the age will be calculated with the documented year of birth, January as imputed month and 1 as imputed day of birth, at the time of informed consent, as follows:

Age (years) = ([Date of Informed Consent] – [Imputed Date of Birth]) / 365.25

It will be summarized as a continuous variable and as a categorical variable based on the following 3 categories: ≤18 years, >18 and <65 years, ≥65 years.

The BMI will be calculated using the following formula (Keys et al, 1972):

$$\text{BMI (kg/m}^2) = 10000 * \frac{\text{weight (kg)}}{[\text{height (cm)}]^2}$$

The BSA for adults is calculated using Mosteller's formula (Mosteller, 1987):

$$\text{BSA (m}^2) = \sqrt{\frac{\text{height (cm)} \times \text{weight (kg)}}{3600}}$$

6.2 Other Baseline characteristics

Disease duration of axSpA will be calculated as the difference between the date of Baseline visit and the onset date of axial spondyloarthritis (axSpA) or Ankylosing spondylitis (AS) as documented on the medical history pages. If the onset date of axSpA is partial, it will be imputed to the most recent feasible date (ie, first day of the month if only day is missing, or the first day of the year if day and month are missing). Disease duration of axSpA will be summarized as continuous variable in years, as well as categorized into <2 years and ≥2 years.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Time since onset of first uveitis flare will be calculated as the difference between the date of Baseline visit and the onset date of first uveitis flare. If the onset date of first uveitis flare is partial, it will be imputed to the most recent feasible date (ie, first day of the month if only day is missing, or the first day of the year if day and month are missing). Time since onset of first uveitis flare will be summarized as continuous variable in years, as well as categorized into <2 years and ≥2 years.

Other baseline characteristics include number of subjects with IBD, number of subjects with psoriasis, number of subjects with at least one tender joint, and number of subjects with at least one swollen joint.

Data collected on childbearing potential will be listed, which includes the method of birth control if the subject is a female of childbearing potential, or if the subject is a fertile male. Also the reason why a male subject is not fertile, or why a female is not of childbearing potential is displayed.

Lifestyle data shows the use of alcohol, caffeinated beverages, illicit drugs, and tobacco. The categories will be summarized on the SS. For subjects with former or current tobacco use, the duration of use is collected and summarized, as well as the number of daily cigarettes, cigars, chewing tobacco, and pipes.

The responses to the following modified ASAS classification criteria for AS will be summarized for all subjects in a frequency table using the SS:

- Back pain of ≥3 months and subject's age at onset <45 years
- Inflammatory back pain
- Reactive arthritis
- Enthesitis
- Uveitis
- Dactylitis
- Psoriasis
- Inflammatory bowel disease
- HLA-B27
- Elevated CRP (above upper limit of normal)
- Sacroiliitis

If medical history terms of back pain or inflammatory back pain are missing for a subject, then the terms axial spondyloarthritis or Ankylosing spondylitis will used instead.

6.3 Medical history and concomitant diseases

The Medical history contains any previous or ongoing medical conditions that occurred prior to study entry. They will be listed for all subjects in the SS including the reported term, start date (month and year only), end date (month and year only) (or ongoing if applicable) and the information whether the procedure is related to axSpA. A glossary of all medical conditions will

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

also be presented including the reported term, the preferred term and the System Organ Class. The classification of medical history will be done according to the MedDRA coding system using system organ class (SOC) and preferred term (PT). These data will be summarized using frequency tables for subject count by classified data.

Procedure history will be displayed in a listing only, data will not be coded.

Extra-articular assessments of psoriasis, IBD, reactive arthritis and uveitis will be presented in listings, based on the subjects in the SS.

6.4 Prior and concomitant medications

Prior medications include any medications that started prior to the start date of study medication. Past medications are a subset of prior medications and include prior medications with a stop date before the date of first study medication administration.

Concomitant medications will be assigned to the treatment period according to whether they were taken at least one day in common with the study medication in that period. That is, if the start or end date of the concomitant medication is on or between the first study drug administration date and the last study drug administration date plus 14 days, or, if the concomitant medication starts prior to first study drug administration and ends after the last study drug administration date plus 14 days.

Concomitant medications will be assigned to the follow-up period if the start date of the concomitant medication is after the last study drug administration date plus 14 days.

A medication can be defined as both prior and concomitant. Where a medication start or stop date is (partially) missing, the medication will be considered a concomitant medication, unless the partial data is suggesting otherwise. Imputation of dates will follow the rules described in section Section 4.2 . Concomitant medications will be summarized for subjects in the SS. The following information will be included in the listings: reported term, dose per intake and dose unit, frequency, formulation, route, indication, start and stop dates (including relative study day), and in case of missing stop date, the information whether the intake of medication is ongoing.

Past medication summaries will be generated for disease-modifying antirheumatic drugs (DMARDs), non-steroidal anti-inflammatory drugs (NSAIDs), and anti-tumor necrosis factors (anti-TNFs) by anatomical therapeutic chemical (ATC) Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT).

Concomitant DMARDs and NSAIDs will be summarized by WHO-DD Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT). Concomitant medications other than DMARDs and NSAIDs will also be summarized.

Intake of any new prohibited medication since last visit is collected and presented in a separate listing.

7 MEASUREMENTS OF TREATMENT COMPLIANCE

There will be 2 approaches to calculate treatment compliance. The first will utilize the number of administered syringes and compare them to the scheduled expected number of syringes. Two syringes should be administered for the loading doses (at weeks 0, 2 and 4) however only 1 syringe should be administered at any other visit from Week 6 until Week 94/last scheduled

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

injection. The difference in number of syringes between the total actual used and the total expected syringes will be summarized. In addition, a ratio of compliance will be further computed based on the number of actual and expected syringes. The ratio of compliance will be summarized as a continuous variable and categorically (<0.80, ≥0.80-≤1.0 and >1.0). The general formula for the compliance ratio (CR) is given as follows:

$$CR = \frac{N_{Syr_act}}{N_{Syr_exp}}$$

where N_{Syr_act} is the number of actual injections and N_{Syr_exp} is the number of expected injections. The second approach defines compliance with study drug administration based upon comparing the actual day of administration with the expected day of administration. The expected day of administration will be based upon the Baseline date and the previous injection date. The sum of the absolute difference in days between the actual and expected days will be summarized. In addition, a ratio of compliance will be further computed based upon the actual and expected day. The ratio of compliance will also be summarized as a continuous variable and categorically (<0.80, ≥0.80-≤1.0 and >1.0). The general formula for the compliance ratio based on duration (CRD) is given as follows:

$$CRD = \frac{T_{Dur_stu} - \Delta_{Cum}}{T_{Dur_stu}}$$

where T_{Dur_stu} is the study duration and Δ_{Cum} is the cumulative difference between actual and scheduled day of administration.

Specific details of this formula are provided in Appendix 12.1 (Compliance Ratio Calculation).

The CRD ranges between 0 and 1. To calculate study duration, the date of the Week 94 visit or the last injection date prior to study treatment discontinuation will be compared to the Baseline date, as shown below:

Study Duration (days) = Week 94 visit/last injection date – Baseline date + 14

The sum will be calculated for the 11 visits from Week 0 (Baseline) to Week 84, and for all planned home administrations from week 6 to week 94. In case the actual day of administration is missing, the maximum deviation to the scheduled day will be assumed (ie, 14 days). If, for a scheduled day with 2 planned injections, the syringes were administered on 2 different days, the maximum day difference of the 2 actual dates to the scheduled date will be utilized.

8 EFFICACY ANALYSES

The efficacy analysis will be based on all subjects in the SS or FAS, as appropriate. If there is no missing AU flare incidence data from pre-study period, the FAS will be the same as the SS. The FAS will be used for analyses of the primary variable and the SS will be used for all other efficacy analyses. The primary efficacy analysis will additionally be repeated for the PPS.

All statistical analyses besides the primary efficacy analysis will be exploratory in nature only. Any p-values or confidence intervals produced in association with secondary or other analyses will be considered non-confirmatory.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

8.1 Statistical analysis of the primary efficacy variable(s)

The primary efficacy analysis will consist of a comparison of the frequency of AU flares during the pre-study/historical period with that observed while in the study during CZP treatment. It is assumed that the frequency of AU flares follows a Poisson distribution. As such, the analysis will be performed as a generalized estimating equations analysis for Poisson outcome that will take into account the possible within-subject correlation (between the retrospective and prospective AU flare counts). The model will contain 2 records per subject corresponding to the frequency of AU flares before study and during the study, offset by log-time of each of the reporting periods. Disease duration of axSpA will be included as a factor in the model, initially defined as <2 years and ≥2 years; the specific cutoffs for analysis may deviate from this based on the distribution seen in the data. The p-value for effect of CZP treatment on frequency of AU flares, along with rate ratio (CZP/historical) and 95% confidence interval, will be obtained from this model. The rate per 96 weeks and rate per 2 years for each study period from the model will also be presented.

The primary efficacy analysis hypothesis will be tested based on significance level of alpha 0.05 and will be 2-sided. The analysis of the other secondary efficacy parameters will be supportive. Therefore, no alpha adjustment for secondary efficacy variables will be performed. The alpha adjustment related to the formal interim review of data is described in Section 4.5.

Statistical Model

Let Y_{ij} be the number of AU flare episodes for subject i in period j, where t_{ij} is the length of the interval. It is assumed that the responses within a subject are correlated. For the statistical model assumptions about the correlation structure need to be done. Assuming an exchangeable correlation structure means that correlation between any two responses of the ith individual is the same. Also we account for different interval length by adding the offset term $\log(t_{ij})$. The mean of the Poisson distribution is $E(Y_{ij}) = \mu_{ij}$, and the mean response per unit time is μ_{ij}/t_{ij}. The Poisson model using a log-link can be written as

$$\log \mu_{ij} = \beta_0 + \beta_1 x_{1;ij} + \beta_2 x_{2;i} + \log t_{ij}$$

which resolves to

$$\mu_{ij} = t_{ij} \exp(\beta_0) * \exp(\beta_1 x_{1;ij}) * \exp(\beta_2 x_{2;i})$$

with the variables:

$x_{1;ij}$: variable for period: 1 on-study, 0 pre-study

$x_{2;i}$: variable for disease duration of axSpA: <2 years, ≥2 years.

It will be tested whether the period variable has an effect on the response variable. The null hypothesis H_0: $\beta_1=0$ of no difference between periods is tested versus the alternative hypothesis H_A: $\beta_1 \neq 0$ of a difference between periods. This 2-sided test will be based on a significance level of alpha 0.05.

8.1.1 Derivations of primary efficacy variable

The primary efficacy variable is the number of distinct episodes of AU flares in the Treatment Period. Per protocol, a flare is considered a new episode if a gap of at least 3 months occurs between 2 flares. If AU flare episodes as reported by the investigator occur less than 3 months

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

(90 days) apart in the same eye, these will be programmatically corrected, such that a new AU flare within 3 months (90 days) after the end of the previous AU flare in the same eye will be considered as pertaining to the same AU episode. These flares will then be counted as one episode. Otherwise, if the time interval between 2 consecutive AU flares is greater than 3 months (90 days), or a separate eye is affected, the AU flares will be considered as separate episodes.

Pre-study and on-study AU flare episodes will be counted. An episode of AU flares is considered as pre-study, if the start date of flare is before the start date of study medication and if start date of flare occurs within the 24 months (730 days) prior to the start of study medication. For the interim analysis, an episode of AU flares is considered as pre-study if the flare occurs within 48 weeks (336 days) prior to the start of study medication.

For the Week 96 (primary) analysis, an episode of AU flares is considered on-study, if the start date of flare is on or after the start date of study medication and the start date of the flare occurs up to and including the last visit in the treatment period, i.e. Week 96 or discontinuation visit.

Pre-study and on-study flares of the same eye will be combined if the time interval between the end of pre-study flare and the start of subsequent on-study flare is less than 3 months (90 days). The resulting flare episode will be considered as pre-study flare episode.

The log-time offset in the Poisson model will consist of the length of pre-study and on-study reporting periods in days. As the documentation of pre-study episodes is extended to up to 2 years prior to Baseline, the pre-study period length for offset is 730 days for each subject, unless the date of first uveitis flare is within 2 years of baseline, in which case the pre-study period length will be from date of first uveitis flare until baseline. The on-study period length for offset is the time in days from start of study medication until last visit, i.e. Week 96 or discontinuation visit. For the interim analysis, an offset of 336 days will be used for the pre-study period or time since first uveitis flare if shorter than 336 days, and the time in days from start of study medication until week 48 visit (taken as 336 days) or early termination visit (taken as 14 days after the last treatment) where applicable will be used for the on-study period.

Disease duration of axSpA, which will be included in the Poisson model as a factor, is calculated as defined in Section 6.2, and is categorized into <2 years and ≥2 years.

8.1.2 Supportive and sensitivity analyses of the primary efficacy variable

The primary efficacy analysis will be repeated in the subgroup of subjects with at least 1 documented AU flare within 12 months prior to Baseline.

The AU flares as reported by the investigators, without programmatically correcting for the 3 month interval between 2 flares, will be analyzed in addition to the primary efficacy variable of protocol-defined AU flares at least 3 months apart.

The primary analysis will be repeated using the pre-study period as 96 weeks (672 days) prior to start of study medication for counting flares and as the offset, instead of 24 months. The on-study period will be the same as that used for the primary analysis.

In addition, the analysis for the full 2 years before and after study medication will be provided. The pre-study period will be the 2 years (730 days) prior to start of study medication, which will be used for counting flares and as the offset. The on-study period will be determined for whose episodes of AU flares for which the start date of flare is on or after the start date of study

This document cannot be used to support any marketing authorisation application or any extensions or variations thereof.
REDACTED COPY

medication and up to and including the SFU visit or date of last administration of study medication + 70 days.

An analysis for the first 48 weeks will also be produced. The period for pre-study will be the 12 months (365 days) prior to start of study medication, which will be used for counting flares and as the offset. An episode of AU flares will be considered on-study, if the start date of flare is on or after the start date of study medication and up to and including the Week 48 visit or discontinuation visit if prior to Week 48. The exact date of the Week 48 or discontinuation visit will be used for each subject. This analysis will also be repeated using 48 weeks (336 days) as the pre-study period for counting flares and as offset, but keeping the on-study period of 48 weeks the same (i.e. using the exact date of the Week 48 or discontinuation visit).

The primary efficacy analysis will be repeated in the subgroup of subjects excluding those with any pre-study flares with completely missing end dates in the 2 years prior to baseline. Subjects will be excluded if any of the flares that started within the 2 years prior to baseline have completely missing end dates (i.e. no day, month or year)

8.2 Statistical analysis of the secondary efficacy variable(s)

Secondary and other efficacy variables will be summarized by time point including change from Baseline or shift from Baseline, with descriptive statistics. Arithmetic mean, SD, median, minimum and maximum value will be displayed for continuous variables. The secondary efficacy variables are assessed at Week 48 and Week 96. All efficacy analyses will be based on the SS or FAS as appropriate.

8.2.1 Number of AU flares per 100 patient-years

The number of AU flares per 100 patient-years in subjects with active axSpA and a history of AU will be calculated and the corresponding 95% confidence interval for both pre-study and on-study periods. Similarly, the number of AU flares per 100 patient-years in subjects with active axSpA and at least 1 AU episode within 12 months prior Baseline, and the corresponding 95% confidence interval will be presented. This variable is also supportive of primary efficacy analysis.

8.2.2 Ankylosing spondylitis disease activity score (ASDAS)

The ASDAS consists of a number of assessments which are scored by the subject and physician and multiplied by a proven formula (van der Heijde et al, 2008) as shown below:

- 0.121 x [redacted] (BASDAI Q2 result)
- 0.058 x [redacted] (BASDAI Q6 result)
- 0.110 x Patient's Global Assessment of Disease Activity (PtGADA)
- 0.073 x [redacted] (BASDAI Q3 result)
- 0.579 x (natural logarithm of the (CRP [mg/L] + 1))

[redacted], PtGADA, [redacted] are all assessed on a numerical scale (0 to 10 units) (Lukas et al, 2009).

The sum of these weighted components gives the ASDAS.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

If 1 component for the ASDAS is missing at a given visit, that component will be imputed by carrying the last observation forward, and the ASDAS will be calculated accordingly. If more than one component for the ASDAS is missing, ASDAS will be treated as missing.

For CRP values below the lower limit of quantification (<4mg/L), half the lower limit (2mg/L) will be used as the imputed value. This approach will be used for summaries of CRP as well.

The nomenclature of disease activity was updated by ASAS (Machado et al, 2018). The state 'Moderate Disease Activity' was replaced by 'Low Disease Activity' state. Disease activity categories based on ASDAS are as follows:

- ASDAS-Inactive Disease (ASDAS-ID): ASDAS <1.3
- ASDAS-Low Disease Activity (ASDAS-LD): ASDAS >=1.3, <2.1
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS >=2.1, <=3.5
- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

The variables related to ASDAS clinical improvement are defined as follows:

- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of >=1.1 relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of >=2.0 relative to Baseline.

Change from Baseline in ASDAS at Week 48 and Week 96 are secondary efficacy variables.

8.2.3 Bath ankylosing spondylitis disease activity index (BASDAI)

The BASDAI is the most commonly used instrument to measure the disease activity of ankylosing spondylitis. The BASDAI is a validated self-reported instrument which consists of 6 horizontal Numeric Rating Scales (NRSs), each with 10 units to measure the severity of the 5 major symptoms: [redacted] over the last week. To give each symptom equal weighting, the average of the 2 scores relating to [redacted] is taken. The resulting 0 to 50 sum score is divided by 5 to give a final BASDAI score between 0 and 10, with lower scores indicating lower disease activity. The BASDAI questions can be found in appendix 16.4 of the clinical study protocol.

If 1 of the 2 [redacted] measurements (ie, questions: [redacted] and [redacted] is missing, the other one will be used for the [redacted] calculation. The same imputation is also applied for the calculation of the ASAS inflammation component, which is calculated as the average of the 2 [redacted] measurements.

If 1 major symptom of the BASDAI is missing, the sum score of the remaining symptoms will be divided by the number of symptoms assessed. If more than 1 major symptom is missing, the sum score will be set to missing.

A response variable called BASDAI50 is defined as an improvement of at least 50% in the BASDAI compared to Baseline.

This document is intended to support a marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Change from Baseline in BASDAI at Week 48 and Week 96 are secondary efficacy endpoints.

8.2.4 ASAS20, ASAS40, and ASAS PR

The ASAS20 response is defined as an improvement of at least 20% and absolute improvement of at least 1 unit on a 0 to 10 NRS in at least 3 of the 4 following domains (Anderson, 2001):

- Patient's Global Assessment of Disease Activity;
- Pain assessment (the total spinal pain NRS score);
- Function (represented by the Bath Ankylosing Spondylitis Functional Index (BASFI));
- Inflammation (the mean of the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) questions [Q] 5 and 6) [redacted])

and absence of deterioration in the potential remaining domain [deterioration is defined as a relative worsening of at least 20% and an absolute worsening of at least 1 unit].

The ASAS criteria for 40% improvement are defined as relative improvement of at least 40% and absolute improvement of at least 2 units on a 0 to 10 NRS in at least 3 of the 4 domains and no worsening at all in the remaining domain.

The ASAS partial remission (PR) response is defined as a score of ≤2 units on a 0 to 10 unit scale in all 4 domains.

8.2.5 Physician's global assessment of disease activity (PhGADA)

The Investigator will assess the overall status of the subject with respect to the axSpA signs and symptoms and the functional capacity of the subject using a visual analog scale where 0 is "very good, asymptomatic and no limitation of normal activities" and 100 is "very poor, very severe symptoms that are intolerable, and the inability to carry out all normal activities." The data including the change from baseline at Week 48 and Week 96 will be summarized as secondary endpoints.

8.2.6 Swollen and tender joint counts

The following 44 joints are to be examined for swelling and tenderness by the Investigator, another delegated physician, or an appropriately qualified medical professional (based on local requirements) who has had documented training on how to perform these assessments correctly. Preferably the same assessor should evaluate the subject at each arthritis assessment.

- Upper body (4) – bilateral sternoclavicular and acromioclavicular joints.
- Upper extremity (26) – bilateral shoulders, elbows, wrists (includes radiocarpal, carpal, and carpometacarpal bones considered as a single unit), metacarpophalangeals (MCPs) I, II, III, IV, and V, and thumb interphalangeals (IPs), and proximal IPs (PIPs) II, III, IV, and V.
- Lower extremity (14) – bilateral knees, ankles, and metatarsophalangeals (I, II, III, IV, and V).

The assessment for swelling and tenderness is made on 44 joints from the above list. Artificial, ankylosed, and missing joints are excluded from swelling and tenderness assessments.

Each joint is scored for tenderness as follows:

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

0 = None (not tender)

1 = Positive (tender)

Each joint is scored for swelling as follows:

0 = None

1 = Detectable

Tender joint counts (TJC) and swollen joint counts (SJC) are calculated as the sum of tender and swollen joints, respectively, among the 44 joints. Both TJC and SJC range from 0 to 44.

If there are missing observations for tender or swollen joints then the remaining observations will be assessed and weighted by dividing by number of non-missing and then multiplying by 44 for the joint count. If a joint is not evaluable at Baseline, then that joint is set to missing throughout the study. If data for more than 50% of the joints are missing at the time of a given assessment, then no imputation will be done and the total TJC or SJC will be set to missing for that visit.

TJC summaries will be based only on those subjects with at least one tender joint at Baseline. Similarly, SJC tables will be based only on those subjects with at least one swollen joint at Baseline.

The data including the change from baseline at Week 48 and Week 96 will be summarized as secondary endpoints.

8.2.7 Patient's global assessment of disease activity (PtGADA)

For the PtGADA questionnaire, subjects will score their global assessment of their disease activity in response to the question "How active was your spondylitis on average during the last week?" using a NRS where 0 is "not active" and 10 is "very active" (van Tubergen et al, 2015). The data including the change from baseline at Week 48 and Week 96 will be summarized as secondary endpoints.

8.2.8 Total and nocturnal spinal pain

The questionnaire for pain in the spine due to AS consists of 2 questions (ie, "How much pain of your spine due to spondylitis do you have?"; and "How much pain of your spine due to spondylitis do you have at night?") (van der Heijde et al, 2005). Total and nocturnal spinal pain will be recorded based on a NRS ranging from 0 to 10, where 0 represents no pain and 10 represents most severe pain. Usually, a 10% difference (ie, a 1 point difference on a NRS ranging from 0 to 10) is considered the minimal clinically important difference used to interpret scores (Dworkin et al, 2008). The total spinal pain data including the change from baseline at Week 48 and Week 96 will be summarized as secondary endpoints.

8.2.9 Bath ankylosing spondylitis functional index (BASFI)

The BASFI contains 10 questions. The first 8 questions [redacted]. The final 2 questions [redacted]. An NRS ranging from 0 to 10 is used to answer the questions on the test. The BASFI questions can be found in appendix 16.7 of the clinical study protocol.

The mean of the 10 scales gives the BASFI score, which is a value between 0 and 10.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

In case of missing answers to 1 or 2 of the single items within the BASFI questionnaire, the BASFI score will be calculated by imputing missing items with the mean of the completed items. Then, the BASFI score will be calculated as described above. If more than 2 of the items are missing, the BASFI score will be left missing.

Change from Baseline in BASFI at Week 48 and Week 96 are secondary efficacy endpoints.

8.3 Analysis of other efficacy variables

All other efficacy endpoints with exception of flare duration and severity will be tabulated at all time points acc to assessment schedule.

8.3.1 Flare assessments

For each subject in FAS, the total duration and the maximum severity of pre-study and on-study AU flares will be determined and summarized. The duration of an episode of AU flare ($T_{\text{Dur_Fl}}$) is calculated as follows:

$$T_{Dur_Fl} = Date_{End_Fl} - Date_{Start_{Fl}} + 1$$

with $Date_{\text{End_Fl}}$ and $Date_{\text{Start_FL}}$ the end date and start date of flare, respectively. The duration of pre-study and on-study periods is defined in Section 4.2.

For the interim analysis, the pre-study period is starting 48 weeks prior to Baseline, i.e. 336 days prior to Baseline until Baseline. The on-study period for the interim analysis is starting from Baseline until week 48, i.e. until study day 336.

For the final analysis, the pre-study period is starting 24 months prior to Baseline, i.e. 730 days prior to Baseline until Baseline. The on-study period for the final analysis is starting from Baseline until week 96, i.e. until the date of the Week 96 or discontinuation visit.

The highest grade (severity) of uveitis flare across all flares per subject will be summarized. The highest grade (severity) noted during an episode of uveitis will be graded as follows: Grade 0=None; Grade 1+=Faint; Grade 2+=Moderate, iris and lens details clear; Grade3+=Marked, iris and lens details hazy; Grade4+=Intense, fibrin or plastic aqueous; ND=Not done or not assessed; UNK=Unknown

8.3.2 Extra-articular assessments

Extra-articular assessments (eg, number of IBD exacerbations and number of psoriasis exacerbations) are to be performed at Screening, Baseline, Weeks 12, 24, 36, 48, 60, 72, 84, and at Completion at Week 96/WD Visit. At each visit, the number of episodes since the last visit and the number of days with IBD/psoriasis are collected. Episodes are considered equivalent to exacerbations. These data will be summarized using summary statistics for continuous variables.

8.3.3 Ankylosing spondylitis quality of life (ASQoL)

The ASQoL consists of 18 items, each with a score of 0 = no or 1 = yes, so that the sum score ranges from 0 to 18, with higher scores indicating worse quality of life. The questions of ASQoL can be found in appendix 16.8 of the clinical study protocol.

If 6 or fewer items are missing, the missing responses will be imputed with the mean of the available responses from that visit to calculate a total score. If more than 6 items are missing, the total score will be left missing.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

The ASQoL total score will be tabulated by visit.

8.3.4 ASAS Health Index (ASAS HI) Questionnaire

The ASAS HI has been developed under the auspices of ASAS to assess health in subjects with all forms of SpA. The questionnaire contains 17 items measuring "functioning, disability and health", these items are given in appendix 16.9 of the clinical study protocol.

Each statement on the ASAS HI is given a score of 1=I agree OR 0=I do not agree. All item scores are summed up to give a total score that ranges from 0 (good functioning) to 17 (poor functioning).

It is to be noted that items 7 and 8 are not applicable for all subjects. For those subjects who ticked the response "not applicable," the sum score is analyzed based on n=16 or n=15, respectively.

Missing data:

A total score can be analyzed if no more than 20% of the data are missing. The total score is calculated as follows for respondents with 1 to a maximum of 3 missing responses:

Sumscore = [x / (17-m)] ·17, where x=Item summation score and m=Number of missing items.

Cases with more than 3 missing responses cannot be allocated a total score. The ASAS HI questionnaire assessments will be listed, and the total score will be tabulated by visit.

8.3.5 Short-Form 36-Item Health Survey (SF-36)

The SF-36 (Version 2, standard recall) is a 36-item generic HRQoL instrument that uses a recall period of 4 weeks. Items are grouped into 8 domains as follows: Physical Functioning (10 items), Role Physical (4 items), Bodily Pain (2 items), General Health (5 items), Vitality (4 items), Social Functioning (2 items), Role Emotional (3 items), Mental Health (5 items), and 1 item for perceived stability or change in health (Health Transition) during the last year. The concepts represented by these domains contribute to physical, mental, and social aspects of HRQoL. The items of SF-36 can be found in appendix 16.10 of the clinical study protocol.

For the calculation of the SF-36 domain scores and the component summaries PCS and MCS, the scoring software QualityMetric Health Outcomes™ Scoring Software 4.5 will be used. The norm-based scores (based on the US general population for 2009) will be utilized for analysis.

9 PHARMACOKINETICS AND PHARMACODYNAMICS

9.1 Pharmacokinetics

Not applicable.

9.2 Pharmacodynamics

Not applicable.

10 SAFETY ANALYSES

All safety analyses will be based on the SS.

10.1 Extent of exposure

The number of doses received will be summarized with continuous statistics for the SS.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Duration of exposure to study medication will be calculated as:

$$Dur_{Exp} = date_{Last} - date_{First} + d_{\text{int}}$$

where d_{int} represents the dosing interval of 14 days.

The cumulative duration of exposure will be summarized for subjects exposed for at least one day, ≥183 days, ≥365 days, ≥548 days, ≥730 days.

Additionally, the subject time at risk will be summarized. Subject time at risk (in days) is calculated as:

$$T_{Risk} = date_{Last} - date_{First} + \left(5 * t_{1/2}\right)$$

where $t_{1/2}$ is 14 days.

If there is a change of treatment or the last clinical contact occurs prior to completion of the follow-up period, the date of last clinical contact per subject is defined as the maximum of [last visit date including safety follow-up visits, last imputed AE start date, last date of termination or completion, last date of study drug administration, death date] or the clinical cutoff date for ongoing studies if earlier than this maximum date).

In cases where the date of last clinical contact (LastCC) occurs prior to completion of this follow-up period, subject time at risk (in days) will be calculated as:

$$T_{Risk} = date_{LastCC} - date_{First} + 1$$

Subject years at risk is the sum of all relevant subject time at risk (in days) divided by 365.25.

The study medication duration will be listed together with the dates of first and last study medication administration.

10.2 Adverse events

Adverse events with start date prior to first administration of study medication are defined as pre-treatment AEs. No summaries of these events will be produced although they will be included in listings.

Adverse events that started more than 70 days after the last administration of study medication are defined as post-treatment AEs. Such events will not be included in tabulated summaries, but will be listed.

Treatment emergent AEs are all AEs starting on or after the date of first study medication and up to 70 days after the last (most recent) dose of study medication. AEs occurring more than 70 days after the last (most recent) dose will not be considered treatment emergent, regardless of whether this occurs during the safety follow-up period, or after a treatment gap of more than 70 days.

If a start date is missing, then the AE will be considered a treatment emergent adverse event (TEAE). Duration should not be calculated if there is missing stop date information.

If the intensity of an adverse event is unknown, it is considered as severe. If the relationship to study drug is missing, it is considered as related.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Tables showing TEAEs, serious TEAEs, non-serious TEAEs, TEAEs leading to discontinuation of study drug, TEAEs by maximum intensity, TEAEs by relationship, serious TEAEs by relationship, fatal TEAEs by relationship, TEAEs including subject numbers will be provided on the SS.

Additionally, more detailed summaries of TEAEs will be presented, which include the exposure adjusted incidence rate (EAIR) with associated 95% confidence interval, and the exposure adjusted event rate (EAER).

The EAIR is defined as the number of subjects (n) with a specific AE adjusted for the exposure and will be scaled to 100 patient-years:

$$EAIR = 100 \cdot n \Big/ \sum_{i=1}^{N} T_{Exp,i}$$

If a subject has multiple events, the time of exposure is calculated to the first occurrence of the AE of interest. If a subject has no events, the total time at risk is used.

Exact Poisson 95% confidence intervals for incidence rates are calculated using the relationship between the Poisson and the Chi-square distribution (Ulm, 1990):

$$LCL = \chi^2_{2n,\alpha/2} \big/ 2$$

$$UCL = \chi^2_{2(n+1),1-\alpha/2} \big/ 2$$

where n is the number of subjects with a specific AE for the incidence rate of interest and is the basis for the number of the degrees of freedom for the chi-square quantile for the upper tail probability χ^2.

The EAER will be the number of AEs including repeat occurrences in individual subjects divided by the total time at risk scaled to 100 patient-years and calculated using:

$$EAER = 100 \cdot N_{AE} \Big/ \sum_{i=1}^{N} T_{Risk,i}$$

where N_{AE} is the total number of AEs.

No confidence interval will be computed for EAER.

Where an AE start date is (partially) missing, the AE will be considered TE if possible. Imputation rules provided in Section 4.2 are based on the assumption that if the start date is incomplete, either day only is missing, or day and months are missing, or date is completely missing. Although the algorithms for treatment-emergence depend on the onset date, imputation rules are provided for resolution date as well, as these may be needed for certain statistical analyses, such as an analyses of AE prevalence or AE duration.

Listings based on SS for all AEs will be provided. A glossary for the reported terms will also be generated utilizing SOC, HLT, and PT.

Anticipated SAEs, as described in Section 10.5 of the clinical study protocol, will not be summarized either by any additional tables or listings, as they are included in the standard AE summaries.

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

AEs of interest will be summarized. The following AEs of interest will be summarized in stand-alone tables:

- Opportunistic infections
- Malignant or unspecified tumors
- Malignant tumors
- Serious cardiovascular events
- Demyelinating-like disorders
- Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia
- Serious bleeding events

For the following AEs of interest, no additional tables will be provided, as these are included in the standard AE summaries:

- Serious infections
- Congestive heart failure
- Lupus and lupus-like illness
- Serious skin reactions (eg, Stevens Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme)

The following notes to AEs of interest in the clinical study protocol are given for reporting purposes only, and therefore will not be summarized separately: A confirmed LTBI must be reported as an AE of interest. Confirmed active TB is an AE of interest. In the event that an active TB occurrence meets the protocol's SAE criteria, it is also an SAE. Potential Hy's Law, defined as ≥3xULN alanine aminotransferase or aspartate aminotransferase with coexisting ≥2xULN total bilirubin in the absence of ≥2xULN alkaline phosphatase, with no alternative explanation for the biochemical abnormality is also considered an AE of interest.

The following events are not officially considered to be AEs of interest but are nonetheless considered to be interesting enough to be summarized in stand-alone tables: 1) Hepatic events and 2) Hypersensitivity reactions and anaphylactic reactions.

The approach for summarizing AEs of interest is detailed in the

UCB AE of interest program specifications for Cimzia from 18 June 2018.

Data collected on hospitalization will be presented in a listing.

10.3 Clinical laboratory evaluations

Hematology, biochemistry, and urinalysis samples will be taken at Screening, Baseline, Weeks 2, 4, 12, 24, 36, 48, 60, 72, 84, and 96/WD Visit. Testing for HBs-Ag, HBc-Ab, HBs-Ab, HCV-Ab, HIV, and HLA-B27 (if not yet known) will be performed at Screening.

The changes from Baseline in laboratory evaluations will be analyzed over time in the SS for observed cases. Shift tables concerning the normal range at each post-baseline visit, will also be produced for each hematology and biochemistry laboratory variable. The shifts will be categorized using L, N, H, missing, and total.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Markedly abnormal laboratory values

The number of subjects with any markedly abnormal hematology or biochemistry value will be tabulated. Values fulfilling the criteria below will be classified as markedly abnormal high (MH) or markedly abnormal low (ML). If no lower (upper) limit is given, the classification ML (MH) is not applicable.

- Hemoglobin < lower limit normal (LLN) and decrease from Baseline >2g/dL
- Hemoglobin <8g/dL
- White blood cells <2000/µL
- Lymphocyte count <1000/µL
- Neutrophil count <1000/µL
- Platelet count <50000/µL
- ALT >3x upper limit normal (ULN)
- AST >3x ULN
- ALP >3x ULN
- Bilirubin ≥2x ULN
- Creatinine >1.8x ULN
- Calcium >12.5mg/dL
- Calcium <7mg/dL
- CK >4x ULN
- Glucose >250mg/dL
- Glucose <40mg/dL
- Potassium >6.4mmol/L
- Potassium <3mmol/L
- Sodium <125mmol/L
- Uric acid ≥3x ULN

The number of subjects meeting the criteria for Hy's Law is of particular importance. The criteria are as follows:

- >=2 x ULN elevation of Bilirubin and 3 x ULN elevation of either ALT or AST

In order to meet the above criteria, a subject must experience the elevation in bilirubin and ALT or AST at the same visit. For example, a subject who experiences a >=2 x ULN elevation of bilirubin at one visit and a 3 x ULN elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's law criteria.

If a repeat sample is taken, the repeated sample data will be used when possible. If the value of a variable collected at the scheduled visit is missing and an additional sample associated with this

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

visit is taken before the next scheduled visit, the missing value will be replaced by this value. Early Withdrawal Visits will be assigned to what would have been the next scheduled visit. In the Urinalysis, microscopy (WBC, RBC, casts, crystals, and bacteria) will be performed only when there are abnormalities on the dipstick. This data will be listed and not presented in tables.

10.4 Vital signs, physical findings, and other observations related to safety

10.4.1 Vital signs

The changes from Baseline in vital signs will be analyzed over time in the SS for observed cases. In addition, the last value (end of treatment) during treatment will be analyzed. End of treatment will be defined as last visit not including the Follow-up visit. Early Withdrawal Visits will be assigned to what would have been the next scheduled visit. The following vital signs measurements will be assessed:

- Pulse rate (bpm)
- Systolic blood pressure (mmHg)
- Diastolic blood pressure (mmHg)
- Body temperature (ºC)

Additionally, body weight (kg) will also be presented together with vital signs.

10.4.2 Pregnancy testing

Pregnancy testing must be carried out for women of childbearing potential and will consist of serum testing at Screening, and urine dipstick testing at Baseline, Week 96/WD Visit, and FU Visit. No tables will be generated and data listed only.

10.4.3 Physical assessments

A physical examination will be performed at Screening, Baseline, Weeks 2, 4, 12, 24, 36, 48, 60, 72, 84, 96/WD Visit, and at the FU Visit. Physical examination findings will be recorded in the eCRF only at Screening. Physical examination data will be listed only.

10.4.4 MRI assessments

Magnetic resonance imaging of the SIJs will be performed at the Screening Visit only. An MRI performed ≤3 months prior to the Baseline Visit may be used as the Baseline. The data will be presented in a listing only.

10.4.5 Tuberculosis assessments

The interferon-gamma release assay (IGRA) test will be done by the central laboratory. The test results will be reported as positive, negative, or indeterminate and must be reviewed by an experienced TB specialist, radiologist, or a pulmonologist. The test result data, the data about history of latent tuberculosis, as well as the TB questionnaire result data, including the 2 questions [redacted] will be presented in listings.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

10.4.6 Chest x-ray

Screening chest x-ray must have occurred within 3 months prior to Screening Visit and should be repeated only if the TB test was confirmed positive or any further evidence is suggestive of potential TB infection. Additional chest x-ray or other imaging test should be performed when positive signs and symptoms indicate pulmonary infection, including potential TB infection, or when close exposure to persons with TB is documented. The data collected on chest x-ray will not be tabulated, but presented in a listing.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

11 REFERENCES

Anderson JJ, Baron G, van der Heijde D, Felson DT, Dougados M. Ankylsoing spondylitis assessment group preliminary definition of short term improvement in ankylosing spondylitis. Arthritis Rheum. 2001 Aug;44(8):1876 86.Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 2008;9:105-21.

Garwood F. Fiducial limits for the Poisson distribution. Biometrika. 1936;28:437-42.

Haybittle JL. Repeated assessments of results in clinical trials of cancer treatment. Brit. J. Radiol. 1971. 44 (526):793-7.

Hu K. Sample Size Estimation for Trials of Recurrent Events. Amgen Inc. 2008. 1-8.

Keys A; Fidanza F; Karvonen M; Kimura N; Taylor H. Indices of relative weight and obesity. Journal of Chronic Diseases 1972, 25 (6–7): 329–43.

Lukas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASASendorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis. 2009; 68(1):18-24.

Machado PM, Landewé R, Heijde Dv, et al. Ann Rheum Dis Epub ahead of print: 16 Feb 2018. Doi:10.1136/annrheumdis-2018-213184

Mosteller RD. Simplified calculation of body surface area. N Engl J Med 1987 Oct22; 317(17):1098 (letter).

Signorini DF. Sample size for Poisson regression. Biometrika. 1991. 78:446-50.

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

van der Heijde D, Sharp J, Wassenberg S, Gladman DD. Psoriatic arthritis imaging: a review of scoring methods. Ann Rheum Dis. 2005;64;ii61-ii64.

van Tubergen A, Black PM, Coteur G. Are patient-reported outcome instruments for ankylosing spondylitis fit for purpose for the axial spondyloarthritis patient? A Qualitative and Psychometric Analysis. Rheumatology. 2015; 54(10):1842-51.

Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary scales: a user's manual. 1994. Maruish, M. E. (Ed.) 2011.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

12 APPENDICES

12.1 APPENDIX 1: Compliance Ratio Calculation

The general formula for treatment compliance is given as follows:

$$CR = \frac{T_{Dur_stu} - \Delta_{Cum}}{T_{Dur_stu}}$$

where T_{Dur_stu} is the study duration and Δ_{Cum} is the cumulative difference between actual and scheduled day of administration. The CRD ranges between 0 and 1.

There are 2 methods to determine the CR. CR (1) is based upon scheduled days which reference the Baseline visits. CR (2) is based upon scheduled days which reference the previous dosing administration date. CR (1) and CR (2) each will be computed separately. The definition of study duration and cumulative difference varies for the 2 calculations. In all cases, if the subject was randomized, but study drug was never administered then the CR is equal to 0. The CR is limited to a range of 0 to 1. Therefore, if for any reason a negative value is computed, the CR must be set to 0. Specific details for each calculation follow.

CR(1):

$$T_{Dur_stu} = Date_{Last} - Date_{BL}$$

$$\Delta_{Cum} = \sum_i |(AD_i - SD_i|$$

where i=Week 0, 2, 4, 6, … , XX

AD_i = Actual day of administration, for Week i=0 to XX by 2
SD_i = (i/2) x 14, for Week i=0 to XX by 2

Note: XX represents the last completed visit in which study drug should have been administered. If subject is still in the study but AD_i is missing, then ABS[AD_i – SD_i]=14 for the missed Visit i.

CR(2):

$$T_{Dur_stu} = Date_{Last} - Date_{BL}$$

$$\Delta_{Cum} = \sum_i |(AD_i - SD_i|$$

where i= Week 0, 2, 4, 6, … , XX

AD_i = Actual day of administration, for Week i=0 to XX by 2
SD_i = {0, for Week i=0 }
{14 + AD_{i-2}, for Week i=2 to XX by 2 }

Note: XX represents the last completed visit in which study drug should have been administered. If subject is still in the study but AD_i is missing, then ABS(AD_i – SD_i)=14 for the missed Visit i.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Example

Week	Scheduled day (1)	Scheduled day (2)	Actual day	ABS Diff (1)	ABS Diff (2)
0	0	0	0	0	0
2	14	14	13	1	1
4	28	27	29	1	2
6	42	43	-	14	14
8	56	57	55	1	2
10	70	69	69	1	0
Total	-	-	-	18	19
CR				66/84 =0.79	65/84 =0.77

Note that this example is for a 12 week study.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

13 AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN (SAP) (IF APPLICABLE)

13.1 AMENDMENT 1

Rationale for the amendment

This SAP was amended after the Data Evaluation Meeting, in order to include the handling of AU flares if the time interval between 2 subsequent flares is less than 3 months. The corresponding derivations for the analysis of the primary efficacy variable, including imputation of start and end dates of flares, have been added. In addition, the interim analysis will include data up to Week 48 only. And according to the most recent AE of interest guidance, separate tables for demyelinating disorders and congestive heart failures were included.

Modifications and changes

Changes in section 4.2

Following has been added at the end of section 4.2:

Efficacy variables

The start and end dates of flares will be imputed for the analyses of AU flares data, in order to allow the calculation of time intervals between 2 consecutive flares and the duration of flares. Different rules will be applied for pre-study and on-study period flares.

Pre-study period (12 months prior to Baseline, i.e. 365 days prior to Baseline, until Baseline):

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Month and year are the same as month and year of the end date of previous flare, and the end date is clearly greater than end date of previous flare, then impute with the end date of previous flare.
  - In the case that month and year of the start date of pre-study period is the same month and year of the end date of previous flare, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Year is the same as year of the end date of previous flare, then impute with the end date of previous flare.
  - In the case that year of the start date of pre-study period is the same year of the end date of previous flare, then use the latest date of both for imputation.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred within 12 months prior to treatment start. Then the maximum (latest) of start date of pre-study period and end date of previous flare will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Month and year are the same as month and year of the start date of subsequent pre-study flare before Baseline, then impute with the start date of subsequent pre-study flare.
  - In the case that month and year of the end date of pre-study period is the same month and year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Year is the same as year of the start date of subsequent pre-study flare, then impute with the start date of subsequent pre-study flare.
  - In the case that year of the end date of pre-study period is the same year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.
- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred within 12 months prior to treatment start. Then the minimum (earliest) of end date of pre-study period and start date of subsequent pre-study flare will be considered.

On-study period (Baseline until end of treatment):

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of study medication, then impute with the start date of study medication.
  - Month and year are the same as month and year of the end date of previous on-study flare, and the end date is clearly greater than end date of previous on-study flare, then impute with the end date of previous on-study flare.
  - In the case that month and year of the start date of study medication is the same month and year of the end date of previous on-study flare, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

  - Year is the same as year of the start date of study medication, then impute with the start date of study medication.
  - Year is the same as year of the end date of previous on-study flare, then impute with the end date of previous on-study flare.
  - In the case that year of the start date of study medication is the same year of the end date of previous on-study flare, then use the latest date of both for imputation.
- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred after treatment start during the on-study period. Then the maximum (latest) of start date of study medication and end date of previous on-study flare will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of on-study period, then impute with the end date of on-study period.
  - Month and year are the same as month and year of the start date of subsequent on-study flare, then impute with the start date of subsequent on-study flare.
  - In the case that month and year of the end date of on-study period is the same month and year of the start date of subsequent on-study flare, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of on-study period, then impute with the end date of on-study period.
  - Year is the same as year of the start date of subsequent on-study flare, then impute with the start date of subsequent on-study flare.
  - In the case that year of the end date of on-study period is the same year of the start date of subsequent on-study flare, then use the earliest date of both for imputation.
- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred after treatment start during the on-study period. Then the minimum (earliest) of end date of on-study period and start date of subsequent on-study flare will be considered.

Changes in section 4.3 (third paragraph)

An interim analysis is planned to be conducted after the last subject has completed the Week 48 assessment. That is when each treated subject has completed the Week 48 Visit, or withdrawn prematurely. All data collected until the database cut-off date will be included. With this interim analysis, first study results will be available already about one year before the final analysis of the study. There is no intention to stop the study early for efficacy or for futility. The interim summaries and listings will be based on the same shells as for the final analysis, unless otherwise stated. The interim listings will include all data collected so far, and also display data after Week

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

48 if available. The summary tables presenting results by visit will only consider results up to Week 48 Visit for the interim analysis. Summary tables of results not provided by visit, e.g. adverse events, concomitant medication, will consist of all the data collected so far at the interim cut-off. All AU flares will be counted for a subject, irrespective of whether a subject has only 48 weeks of data or already more weeks of data collected.

Has been changed to:

An interim analysis is planned to be conducted after the last subject has completed the Week 48 assessment. That is when each treated subject has completed the Week 48 Visit, or withdrawn prematurely. All data collected until the database cut-off date will be included. With this interim analysis, first study results will be available already about one year before the final analysis of the study. There is no intention to stop the study early for efficacy or for futility. The interim summaries and listings will be based on the same shells as for the final analysis, unless otherwise stated. The interim listings will include all data collected **up to** Week 48. The summary tables presenting results by visit will only consider results up to Week 48 Visit for the interim analysis. Summary tables of results not provided by visit, e.g. adverse events, concomitant medication, will **not include events or medication that started after the Week 48 Visit. Data collected after Week 48 will not be presented in the interim analysis.**

Changes in section 6.2 (second paragraph)

Time since onset of first uveitis flare will be calculated as the difference between the date of Baseline visit and the onset date of first uveitis flare. If the onset date of first uveitis flare is partial, it will be imputed to the most recent feasible date (ie, last day of the month if only day is missing, or the last day of the year if day and month are missing). Time since onset of first uveitis flare will be summarized as continuous variable in days, as well as categorized into <2 years and ≥2 years.

Has been changed to:

Time since onset of first uveitis flare will be calculated as the difference between the date of Baseline visit and the onset date of first uveitis flare. If the onset date of first uveitis flare is partial, it will be imputed to the **first** feasible date (ie, **first** day of the month if only day is missing, or the **first** day of the year if day and month are missing). Time since onset of first uveitis flare will be summarized as continuous variable in days, as well as categorized into <2 years and ≥2 years.

Changes in section 8.1.1

The primary efficacy variable is the number of distinct episodes of AU flares in the Treatment Period. Pre-study and on-study AU flare episodes will be counted. An episode of AU flares is considered as pre-study, if the start date of flare is before the start date of study medication. An episode of AU flares is considered on-study, if the start date of flare is on or after the start date of study medication.

The log-time offset in the Poisson model will consist of the length of pre-study and on-study reporting periods in days. As the documentation of pre-study episodes starts 12 months prior to Baseline, the pre-study period length for offset is 365 days for each subject. The on-study period length for offset is the time in days from start of study medication until last visit.

This document cannot be used to support any marketing application and any extensions or variations thereof.
REDACTED COPY

Disease duration of axSpA, which will be included in the Poisson model as a factor, is calculated from the start date of first AU flare episode until the Baseline visit date, and is categorized into <2 years and ≥2 years.

Has been changed to:

The primary efficacy variable is the number of distinct episodes of AU flares in the Treatment Period. **Per protocol, a flare is considered a new episode if a gap of at least 3 months occurs between 2 flares. If AU flare episodes as reported by the investigator occur less than 3 months (90 days) apart in the same eye, these will programmatically corrected, such that a new AU flare within 3 months (90 days) after the end of the previous AU flare in the same eye will be considered as pertaining to the same AU episode. These flares will then be counted as one episode. Otherwise, if the time interval between 2 consecutive AU flares is greater than 3 months (90 days), or a separate eye is affected, the AU flares will be considered as separate episodes.**

Pre-study and on-study AU flare episodes will be counted. An episode of AU flares is considered as pre-study, if the start date of flare is before the start date of study medication **and if the flare occurs within the 12 months (365 days) prior to the start of study medication**. An episode of AU flares is considered on-study, if the start date of flare is on or after the start date of study medication. **Pre-study and on-study flares of the same eye will be combined if the time interval between the end of pre-study flare and the start of subsequent on-study flare is less than 3 months (90 days). The resulting flare episode will be considered as pre-study flare episode.**

The log-time offset in the Poisson model will consist of the length of pre-study and on-study reporting periods in days. As the documentation of pre-study episodes starts 12 months prior to Baseline, the pre-study period length for offset is 365 days for each subject. The on-study period length for offset is the time in days from start of study medication until last visit.

Disease duration of axSpA, which will be included in the Poisson model as a factor, is calculated from the start date of first AU flare episode until the Baseline visit date, and is categorized into <2 years and ≥2 years.

Changes in section 8.1.2

The primary efficacy analysis will be repeated in the subgroup of subjects with at least 1 documented AU flare within 12 months prior to Baseline.

The AU flare rate (event rate) per 100 patient-years of exposure, both before and during the study will be calculated. The corresponding rate ratio along with 95% confidence interval (Garwood, 1936) will be presented as a key supportive analysis to the primary efficacy analysis results (see Section 8.2.1).

Has been changed to:

The primary efficacy analysis will be repeated in the subgroup of subjects with at least 1 documented AU flare within 12 months prior to Baseline.

The AU flares as reported by the investigators, without programmatically correcting for the 3 month interval between 2 flares, will be analyzed in addition to the primary efficacy variable of protocol-defined AU flares at least 3 months apart.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

The AU flare rate (event rate) per 100 patient-years of exposure, both before and during the study will be calculated. The corresponding rate ratio along with 95% confidence interval (Garwood, 1936) will be presented as a key supportive analysis to the primary efficacy analysis results (see Section 8.2.1).

Changes in section 8.2.2 (last paragraph)

The variables related to ASDAS clinical improvement are defined as follows:

- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of >=1.1 relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of >=2.0 relative to Baseline

Change from Baseline in ASDAS at Week 48 and Week 96 are secondary efficacy variables.

Has been changed to:

The variables related to ASDAS clinical improvement are defined as follows:

- **ASDAS- Clinical Improvement (ASDAS-CI): ASDAS reduction (improvement) of >0 relative to Baseline**
- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of >=1.1 relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of >=2.0 relative to Baseline **or has the lowest score possible post-baseline (i.e. when CRP<LLOQ and all other components are 0, then the minimum ASDAS score is 0.636 to 3 decimal places).**

Change from Baseline in ASDAS at Week 48 and Week 96 are secondary efficacy variables.

Changes in section 10.2

AEs of interest will be summarized. The following AEs of interest will be summarized in stand-alone tables:

- Opportunistic infections
- Malignant and unspecified tumors
- Malignant tumors
- Serious cardiovascular events
- Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia
- Serious bleeding events

For the following AEs of interest, no additional tables will be provided, as these are included in the standard AE summaries:

- Serious infections
- Demyelinating-like disorders

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- Congestive heart failure
- Lupus and lupus-like illness
- Serious skin reactions (eg, Stevens Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme)

The following notes to AEs of interest in the clinical study protocol are given for reporting purposes only, and therefore will not be summarized separately: A confirmed LTBI must be reported as an AE of interest. Confirmed active TB is an AE of interest. In the event that an active TB occurrence meets the protocol's SAE criteria, it is also an SAE. Potential Hy's Law, defined as ≥3xULN alanine aminotransferase or aspartate aminotransferase with coexisting ≥2xULN total bilirubin in the absence of ≥2xULN alkaline phosphatase, with no alternative explanation for the biochemical abnormality.

Has been changed to:

AEs of interest will be summarized. The following AEs of interest will be summarized in stand-alone tables:

- Opportunistic infections
- Malignant and unspecified tumors
- Malignant tumors
- Serious cardiovascular events
- **Demyelinating-like disorders**
- Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia
- Serious bleeding events

For the following AEs of interest, no additional tables will be provided, as these are included in the standard AE summaries:

- Serious infections
- Congestive heart failure
- Lupus and lupus-like illness
- Serious skin reactions (eg, Stevens Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme)

The following notes to AEs of interest in the clinical study protocol are given for reporting purposes only, and therefore will not be summarized separately: A confirmed LTBI must be reported as an AE of interest. Confirmed active TB is an AE of interest. In the event that an active TB occurrence meets the protocol's SAE criteria, it is also an SAE. Potential Hy's Law, defined as ≥3xULN alanine aminotransferase or aspartate aminotransferase with coexisting ≥2xULN total bilirubin in the absence of ≥2xULN alkaline phosphatase, with no alternative explanation for the biochemical abnormality **is also considered an AE of interest.**

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

13.2 AMENDMENT 2

Rationale for the amendment

This SAP was modified to implement decisions from the second Data Evaluation Meeting. This includes the updated definition of Baseline to consider also assessments that were done shortly after the first study medication administration but still on the same day of first study medication administration. In addition, the SAP was revised to include collection of ASAS classification criteria as part of the medical history following the extension of the pre-study period for AU flare collection.

Modifications and changes

Changes in section 3.3

Unless otherwise specified, the last valid measurement before study medication administration will be utilized as Baseline value. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables, assessments may be scheduled for Screening only and not for Baseline. In this case the screening value will be utilized as baseline value. If a baseline measurement is missing, and a screening value available, the screening value will be utilized as Baseline instead. If a Baseline visit measurement is missing, and a Screening visit measurement is available, the Screening value will be utilized as Baseline. If a scheduled Baseline assessment is taken on the same day as first administration of study medication and no time is collected, it will be assumed to have been taken prior to study medication.

Has been changed to:

Unless otherwise specified, the last valid measurement before study medication administration will be utilized as Baseline value. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables, assessments may be scheduled for Screening only and not for Baseline. In this case the screening value will be utilized as baseline value. If a baseline measurement is missing, and a screening value available, the screening value will be utilized as Baseline instead. If a Baseline visit measurement is missing, and a Screening visit measurement is available, the Screening value will be utilized as Baseline. **If a scheduled Baseline assessment is taken on the same day as first administration of study medication, it will be used as Baseline, irrespective of the time of the assessment relative to the first administration of study medication.**

Changes in section 3.5.3

The FAS will consist of all subjects in the SS with non-missing Baseline value for the primary efficacy variable, i.e. AU flare incidence data from pre-study period.

In the unlikely event that the FAS differs from the SS, the efficacy analysis will be conducted using the FAS instead of the SS.

Has been changed to:

The FAS will consist of all subjects in the SS with non-missing Baseline value for the primary efficacy variable, i.e. AU flare incidence data from pre-study period.

If the FAS differs from the SS, the efficacy analysis will be conducted using the FAS instead of the SS.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Changes in section 4.2 (Efficacy variables)

The start and end dates of flares will be imputed for the analyses of AU flares data, in order to allow the calculation of time intervals between 2 consecutive flares and the duration of flares. Different rules will be applied for pre-study and on-study period flares.

Pre-study period (12 months prior to Baseline, i.e. 365 days prior to Baseline, until Baseline):

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Month and year are the same as month and year of the end date of previous flare, and the end date is clearly greater than end date of previous flare, then impute with the end date of previous flare.
  - In the case that month and year of the start date of pre-study period is the same month and year of the end date of previous flare, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Year is the same as year of the end date of previous flare, then impute with the end date of previous flare.
  - In the case that year of the start date of pre-study period is the same year of the end date of previous flare, then use the latest date of both for imputation.
- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred within 12 months prior to treatment start. Then the maximum (latest) of start date of pre-study period and end date of previous flare will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Month and year are the same as month and year of the start date of subsequent pre-study flare before Baseline, then impute with the start date of subsequent pre-study flare.
  - In the case that month and year of the end date of pre-study period is the same month and year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of pre-study period, then impute with the end date of pre-study period.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- Year is the same as year of the start date of subsequent pre-study flare, then impute with the start date of subsequent pre-study flare.
- In the case that year of the end date of pre-study period is the same year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.

Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred within 12 months prior to treatment start. Then the minimum (earliest) of end date of pre-study period and start date of subsequent pre-study flare will be considered.

Has been changed to:

The start and end dates of flares will be imputed for the analyses of AU flares data, in order to allow the calculation of time intervals between 2 consecutive flares and the duration of flares. Different rules will be applied for pre-study and on-study period flares.

For the interim analysis, the pre-study period is starting 48 weeks prior to Baseline, i.e. 336 days prior to Baseline until Baseline. The on-study period for the interim analysis is starting from Baseline until week 48, i.e. until study day 336.

For the final analysis, the pre-study period is starting 96 weeks prior to Baseline, i.e. 672 days prior to Baseline until Baseline. The on-study period for the final analysis is starting from Baseline until week 96, i.e. until study day 672. In addition, the period definition used for interim with period length of 48 weeks will be repeated for the final analysis.

Pre-study period (**48 vs 96 weeks** prior to Baseline, i.e. **336 vs 672** days prior to Baseline, until Baseline **for interim vs final analysis**):

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Month and year are the same as month and year of the end date of previous flare, and the end date is clearly greater than end date of previous flare, then impute with the end date of previous flare.
  - In the case that month and year of the start date of pre-study period is the same month and year of the end date of previous flare, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Year is the same as year of the end date of previous flare, then impute with the end date of previous flare.
  - In the case that year of the start date of pre-study period is the same year of the end date of previous flare, then use the latest date of both for imputation.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred within **the pre-study period**. Then the maximum (latest) of start date of pre-study period and end date of previous flare will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Month and year are the same as month and year of the start date of subsequent pre-study flare before Baseline, then impute with the start date of subsequent pre-study flare.
  - In the case that month and year of the end date of pre-study period is the same month and year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Year is the same as year of the start date of subsequent pre-study flare, then impute with the start date of subsequent pre-study flare.
  - In the case that year of the end date of pre-study period is the same year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.
- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred within **the pre-study period**. Then the minimum (earliest) of end date of pre-study period and start date of subsequent pre-study flare will be considered.

Changes in section 4.3

Regular monitoring of safety data collected during clinical studies will be performed as described in the Safety Signal Detection in the Ongoing Clinical Trials Charter for CZP.

A specific data monitoring, steering, or evaluation committee is not planned for this study.

An interim analysis is planned to be conducted after the last subject has completed the Week 48 assessment. That is when each treated subject has completed the Week 48 Visit, or withdrawn prematurely. All data collected until the database cut-off date will be included. With this interim analysis, first study results will be available already about one year before the final analysis of the study. There is no intention to stop the study early for efficacy or for futility. The interim summaries and listings will be based on the same shells as for the final analysis, unless otherwise stated. The interim listings will include all data collected up to Week 48. The summary tables presenting results by visit will only consider results up to Week 48 Visit for the interim analysis. Summary tables of results not provided by visit, e.g. adverse events, concomitant medication, will not include events or medication that started after the Week 48 Visit. Data collected after Week 48 will not be presented in the interim analysis.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

No separate SAP will be created for the interim analyses, as all analyses will be covered within this SAP, and within the corresponding display specifications. In general, the interim analysis will follow the same specifications as the final analysis, unless stated differently in the corresponding sections. The interim analysis will include a subset of tables and listings that will be marked in the display specifications.

An additional interim data-cut might be needed to support a regulatory submission in nr-axSpA.

Has been changed to:

Regular monitoring of safety data collected during clinical studies will be performed as described in the Safety Signal Detection in the Ongoing Clinical Trials Charter for CZP.

A specific data monitoring, steering, or evaluation committee is not planned for this study.

An interim analysis is planned to be conducted after the last subject has completed the Week 48 assessment. That is when each treated subject has completed the Week 48 Visit, or withdrawn prematurely. All data collected until the database cut-off date will be included. With this interim analysis, first study results will be available already about one year before the final analysis of the study. There is no intention to stop the study early for efficacy or for futility. The interim summaries and listings will be based on the same shells as for the final analysis, unless otherwise stated. The interim listings will include all data collected up to Week 48. The summary tables presenting results by visit will only consider results up to Week 48 Visit for the interim analysis. Summary tables of results not provided by visit, e.g. adverse events, will not include events or medication that started after the Week 48 Visit. Data collected after Week 48 will not be presented in the interim analysis.

The length of pre-study period and on-study period for collection of AU flares data will be 48 weeks for the interim analysis, and will be 96 weeks for the final analysis. The analysis based on period length of 48 weeks will be repeated for the final analysis.

No separate SAP will be created for the interim analyses, as all analyses will be covered within this SAP, and within the corresponding display specifications. In general, the interim analysis will follow the same specifications as the final analysis, unless stated differently in the corresponding sections. The interim analysis will include a subset of tables and listings that will be marked in the display specifications.

An additional interim data-cut might be needed to support a regulatory submission in nr-axSpA.

Changes in section 6.2

The following sentence has been added at the end of section 6.2:

The responses to the following modified ASAS classification criteria for AS will be summarized for all subjects in a frequency table using the SS:

- **Back pain of ≥3 months and subject's age at onset <45 years**
- **Inflammatory back pain**
- **Reactive arthritis**
- **Enthesitis**
- **Uveitis**

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- **Dactylitis**
- **Psoriasis**
- **Inflammatory bowel disease**
- **HLA-B27**
- **Elevated CRP (above upper limit of normal)**
- **Sacroiliitis**

Changes in section 8.1 (second paragraph)

The primary efficacy analysis hypothesis will be tested based on significance level of alpha 0.05 and will be 2-sided. The entire alpha will be used for the primary efficacy analysis and analysis of the other secondary efficacy parameters will be supportive. Therefore, no alpha adjustment will be performed.

Has been changed to:

The primary efficacy analysis hypothesis will be tested based on significance level of alpha 0.05 and will be 2-sided. The analysis of the other secondary efficacy parameters will be supportive. Therefore, no alpha adjustment **for secondary efficacy variables** will be performed. **The alpha adjustment related to the formal interim review of data is described in Section 4.5.**

Changes in section 8.1.1

The primary efficacy variable is the number of distinct episodes of AU flares in the Treatment Period. Per protocol, a flare is considered a new episode if a gap of at least 3 months occurs between 2 flares. If AU flare episodes as reported by the investigator occur less than 3 months (90 days) apart in the same eye, these will be programmatically corrected, such that a new AU flare within 3 months (90 days) after the end of the previous AU flare in the same eye will be considered as pertaining to the same AU episode. These flares will then be counted as one episode. Otherwise, if the time interval between 2 consecutive AU flares is greater than 3 months (90 days), or a separate eye is affected, the AU flares will be considered as separate episodes.

Pre-study and on-study AU flare episodes will be counted. An episode of AU flares is considered as pre-study, if the start date of flare is before the start date of study medication and if the flare occurs within the 12 months (365 days) prior to the start of study medication. An episode of AU flares is considered on-study, if the start date of flare is on or after the start date of study medication. Pre-study and on-study flares of the same eye will be combined if the time interval between the end of pre-study flare and the start of subsequent on-study flare is less than 3 months (90 days). The resulting flare episode will be considered as pre-study flare episode.

The log-time offset in the Poisson model will consist of the length of pre-study and on-study reporting periods in days. As the documentation of pre-study episodes starts 12 months prior to Baseline, the pre-study period length for offset is 365 days for each subject. The on-study period length for offset is the time in days from start of study medication until last visit.

Disease duration of axSpA, which will be included in the Poisson model as a factor, is calculated from the start date of first AU flare episode until the Baseline visit date, and is categorized into <2 years and ≥2 years.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Has been changed to:

The primary efficacy variable is the number of distinct episodes of AU flares in the Treatment Period. Per protocol, a flare is considered a new episode if a gap of at least 3 months occurs between 2 flares. If AU flare episodes as reported by the investigator occur less than 3 months (90 days) apart in the same eye, these will be programmatically corrected, such that a new AU flare within 3 months (90 days) after the end of the previous AU flare in the same eye will be considered as pertaining to the same AU episode. These flares will then be counted as one episode. Otherwise, if the time interval between 2 consecutive AU flares is greater than 3 months (90 days), or a separate eye is affected, the AU flares will be considered as separate episodes.

Pre-study and on-study AU flare episodes will be counted. An episode of AU flares is considered as pre-study, if the start date of flare is before the start date of study medication and if the flare occurs within the **96 weeks** (**672** days) prior to the start of study medication. **For the interim analysis, an episode of AU flares is considered as pre-study if the flare occurs within 48 weeks (336 days) prior to the start of study medication.** An episode of AU flares is considered on-study, if the start date of flare is on or after the start date of study medication. Pre-study and on-study flares of the same eye will be combined if the time interval between the end of pre-study flare and the start of subsequent on-study flare is less than 3 months (90 days). The resulting flare episode will be considered as pre-study flare episode.

The log-time offset in the Poisson model will consist of the length of pre-study and on-study reporting periods in days. As the documentation of pre-study episodes **is extended to up to 2 years** prior to Baseline, the pre-study period length for offset is **672** days for each subject. The on-study period length for offset is the time in days from start of study medication until last visit. **For the interim analysis, an offset of 336 days will be used for the pre-study period, and the time in days from start of study medication until week 48 visit or early termination visit where applicable will be used.**

Disease duration of axSpA, which will be included in the Poisson model as a factor, is calculated from the start date of first AU flare episode until the Baseline visit date, and is categorized into <2 years and ≥2 years.

Changes in section 8.2.2

The ASDAS consists of a number of assessments which are scored by the subject and physician and multiplied by a proven formula (van der Heijde et al, 2008) as shown below:

- 0.121 x [redacted] (BASDAI Q2 result)
- 0.058 x [redacted] (BASDAI Q6 result)
- 0.110 x Patient's Global Assessment of Disease Activity (PtGADA)
- 0.073 x [redacted] (BASDAI Q3 result)
- 0.579 x (natural logarithm of the (CRP [mg/L] + 1))

[redacted], PtGADA, [redacted] are all assessed on a numerical scale (0 to 10 units) (Lukas et al, 2009).

The sum of these weighted components gives the ASDAS.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

If 1 component for the ASDAS is missing at a given visit, that component will be imputed by carrying the last observation forward, and the ASDAS will be calculated accordingly. If more than one component for the ASDAS is missing, ASDAS will be treated as missing.

For CRP values below the lower limit of quantification (<4mg/L), half the lower limit (2mg/L) will be used as the imputed value.

Disease activity categories based on ASDAS are as follows:

- ASDAS-Inactive Disease (ASDAS-ID): ASDAS <1.3
- ASDAS-**Moderate Disease (ASDAS-MD)**: ASDAS >=1.3, <2.1
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS >=2.1, <=3.5
- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

The variables related to ASDAS clinical improvement are defined as follows:

- ASDAS-Clinical Improvement (ASDAS-CI): ASDAS reduction (improvement) of >0 relative to Baseline
- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of >=1.1 relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of >=2.0 relative to Baseline or has the lowest score possible post-baseline (i.e. when CRP<LLOQ and all other components are 0, then the minimum ASDAS score is 0.636 to 3 decimal places).

Change from Baseline in ASDAS at Week 48 and Week 96 are secondary efficacy variables.

Has been changed to:

The ASDAS consists of a number of assessments which are scored by the subject and physician and multiplied by a proven formula (van der Heijde et al, 2008) as shown below:

- 0.121 x [redacted] (BASDAI Q2 result)
- 0.058 x [redacted] (BASDAI Q6 result)
- 0.110 x Patient's Global Assessment of Disease Activity (PtGADA)
- 0.073 x [redacted] (BASDAI Q3 result)
- 0.579 x (natural logarithm of the (CRP [mg/L] + 1))

[redacted] PtGADA, [redacted] are all assessed on a numerical scale (0 to 10 units) (Lukas et al, 2009).

The sum of these weighted components gives the ASDAS.

If 1 component for the ASDAS is missing at a given visit, that component will be imputed by carrying the last observation forward, and the ASDAS will be calculated accordingly. If more than one component for the ASDAS is missing, ASDAS will be treated as missing.

For CRP values below the lower limit of quantification (<4mg/L), half the lower limit (2mg/L) will be used as the imputed value.

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

The nomenclature of disease activity was updated by ASAS (Machado et al, 2018). The state 'Moderate Disease Activity' was replaced by 'Low Disease Activity' state. Disease activity categories based on ASDAS are as follows:

- ASDAS-Inactive Disease (ASDAS-ID): ASDAS <1.3
- ASDAS-**Low Disease Activity (ASDAS-LD)**: ASDAS >=1.3, <2.1
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS >=2.1, <=3.5
- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

The variables related to ASDAS clinical improvement are defined as follows:

- ASDAS-Clinical Improvement (ASDAS-CI): ASDAS reduction (improvement) of >0 relative to Baseline
- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of >=1.1 relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of >=2.0 relative to Baseline or has the lowest score possible post-baseline (i.e. when CRP<LLOQ and all other components are 0, then the minimum ASDAS score is 0.636 to 3 decimal places).

Change from Baseline in ASDAS at Week 48 and Week 96 are secondary efficacy variables.

Changes in section 11

The following reference was added:

Machado PM, Landewé R, Heijde Dv, et al. Ann Rheum Dis Epub ahead of print: 16 Feb 2018. Doi:10.1136/annrheumdis-2018-213184

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

13.3 AMENDMENT 3

Rationale for the amendment

This SAP was modified to implement decisions that were made before the fourth Data Evaluation Meeting. This includes the updated definitions of pre-study and on-study periods for the primary analysis to consider also AU flare episodes that occurred within 24 months (730 days) prior to the start study medication and starting from Baseline until Week 96 or discontinuation visit and for supportive and sensitivity analysis to consider pre-study period as 96 weeks (672 days) prior to start of study medication for counting flares and as the offset. In addition, the SAP was revised to make some minor administrative changes throughout.

Modifications and changes

Changes in section 1

The SAP is based on **Amendment 2** of the Clinical Study Protocol, dated **17 Jan 2017**.

Has been changed to:

The SAP is based on **Amendment 3** of the Clinical Study Protocol, dated **22 Jan 2020**.

Add section 2.1.4

The safety objectives of the study will be to evaluate the safety and tolerability of CZP therapy.

Changes in section 2.2.2

Safety variables to be assessed are:

- Adverse events (AEs)
- Blood pressure
- Physical examination
- Clinical laboratory values (hematology, biochemistry, and urinalysis)

Has been changed to:

2.2.2.1 Secondary safety variable

The secondary safety variable to be assessed is treatment-emergent adverse events (TEAEs).

2.2.2.2 Other safety variables

The following other safety variables to be assessed are:

- Blood pressure
- Physical examination
- Clinical laboratory values (hematology, biochemistry, and urinalysis)

Changes in section 3.1

Unless stated otherwise, all inferential statistical tests will be 2-sided and conducted at the 0.05 alpha level. Adjustment for multiplicity is described in Section 4.5. P-values for secondary efficacy variables are nominal. P-values will be presented to 3 decimal places.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Has been changed to:

Unless stated otherwise, all inferential statistical tests will be 2-sided and conducted at the 0.05 alpha level. Adjustment for multiplicity is described in Section 4.5. P-values for secondary efficacy variables are nominal. P-values will be presented to 3 decimal **places and any values below 0.001 should be displayed as "<0.001".**

Changes in section 3.3

Unless otherwise specified, the last valid measurement before study medication administration will be utilized as Baseline value. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables, assessments may be scheduled for Screening only and not for Baseline. In this case the screening value will be utilized as baseline value. If a baseline measurement is missing, and a screening value available, the screening value will be utilized as Baseline instead. **If a Baseline visit measurement is missing, and a Screening visit measurement is available, the Screening value will be utilized as Baseline.** If a scheduled Baseline assessment is taken on the same day as first administration of study medication, it will be used as Baseline, irrespective of the time of the assessment relative to the first administration of study medication.

Has been changed to:

Unless otherwise specified, the last valid measurement before study medication administration will be utilized as Baseline value. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables, assessments may be scheduled for Screening only and not for Baseline. In this case the screening value will be utilized as baseline value. If a baseline measurement is missing, and a screening value available, the screening value will be utilized as Baseline instead. If a scheduled Baseline assessment is taken on the same day as first administration of study medication, it will be used as Baseline, irrespective of the time of the assessment relative to the first administration of study medication

Changes in section 3.4

Important protocol deviations (IPD) are deviations from the protocol which could potentially have a meaningful impact on study conduct, on the safety for an individual subject, or on the primary efficacy outcome for an individual subject. The criteria for identifying important protocol deviations will be defined within the specification of the IPD document. Important protocol deviations will be reviewed as part of the ongoing data cleaning process and all important deviations will be identified and documented prior to database lock, **although no analysis will be performed based on per protocol population.**

Important protocol deviations will be identified and classified by the deviation types defined in the IPD document. A listing of all important protocol deviations identified at the Data Evaluation Meeting will be presented for all subjects in the SS, and will include deviation type and description. The number and percentage of subjects in the **ES** with important protocol deviations will be summarized overall **and by treatment**.

Has been changed to:

Important protocol deviations (IPD) are deviations from the protocol which could potentially have a meaningful impact on study conduct, on the safety for an individual subject, or on the primary efficacy outcome for an individual subject. The criteria for identifying important

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

protocol deviations will be defined within the specification of the IPD document. Important protocol deviations will be reviewed as part of the ongoing data cleaning process and all important deviations will be identified and documented prior to database lock**.**

Important protocol deviations will be identified and classified by the deviation types defined in the IPD document. A listing of all important protocol deviations identified at the Data Evaluation Meeting will be presented for all subjects in the SS, and will include deviation type and description. The number and percentage of subjects in the **SS** with important protocol deviations will be summarized overall.

Changes in section 3.5.3

If the FAS differs from the SS, the efficacy analysis will be conducted using the FAS instead of the SS.

Has been changed to:

In case the FAS differs from the SS, the efficacy analysis will be conducted using the FAS instead of the SS.

Changes in section 3.9

As the BASMI is not collected, and there is no alternative measure of spinal mobility available in the study data, the complete component for spinal mobility is missing. Therefore the ASAS 5/6 response criterion cannot be calculated, and the analysis of the secondary efficacy variable ASAS 5/6 has to be dropped.

The clinical study protocol does not mention a PPS. This analysis set has been added, in order to additionally perform the primary efficacy analysis on the subset of protocol-compliant subjects.

ASDAS Clinical improvement was detailed as an "other" efficacy variable, but no definition was given. As there is no well-accepted definition, this variable has not been included in the analysis.

Has been changed to:

As the BASMI is not collected, and there is no alternative measure of spinal mobility available in the study data, the complete component for spinal mobility is missing. Therefore the ASAS 5/6 response criterion cannot be calculated, and the analysis of the secondary efficacy variable ASAS 5/6 has to be dropped.

The clinical study protocol does not mention a PPS. This analysis set has been added, in order to additionally perform the primary efficacy analysis on the subset of protocol-compliant subjects.

In the calculation of ASDAS (Section 9.1.2 of the protocol), the CRP component was written as "0.579 × (natural logarithm of the CRP [mg/L] + 1)", but should have been "0.579 × (natural logarithm of the (CRP [mg/L] + 1))".

For BASDAI (Section 9.1.3 of the protocol), the protocol stated "The BASDAI is a validated self-reported instrument which consists of six 10-unit horizontal NRSs to

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

[redacted]

In Section 9.1.4 of the protocol, the pain assessment component of ASAS20/40/PR was defined as "the average of total and nocturnal spinal pain NRS scores", but this should be defined as total spinal pain NRS score only.

ASDAS Clinical improvement was detailed as an "other" efficacy variable, but no definition was given. As there is no well-accepted definition, this variable has not been included in the analysis.

In Section 12.7 of the protocol, it stated that the interim analysis would use α=0.0001 and so the final analysis would use α=0.049. This should have said that α=0.001 would be used for the interim analysis.

Changes in section 4.2

Analysis based on observed case **and last observation carried forward for continuous variables or non-responder imputation for categorical variables will be presented**.

If the Baseline value is missing, the Screening value may be used as the Baseline. In general, if no measurement is available prior to receiving study medication, then the Baseline value is treated as missing, and will not be imputed.

Post-Baseline values may only be imputed by carrying forward earlier post-Baseline values. Baseline values will not be carried forward to post-Baseline time points.

Premature withdrawal visits should be assigned to what would have been the next scheduled visit for the particular assessment that is being summarized. If a premature withdrawal visit occurs on the same date as a scheduled visit, then the assessments will be summarized with the scheduled visit.

Safety variables

For analyses of AEs and concomitant medication usage, a complete date must be established in order to correctly identify the AE or medication as occurring during treatment or not. For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the subject data listings (i.e., no imputed values will be displayed in data listings).

- Missing start day, but month and year present: If the start of study medication occurred in the same month and year as the occurrence of the AE/concomitant medication, the start day of the event/concomitant medication will be assigned to the day of first intake of study medication. Otherwise the start day will be set to the 1st day of the month.
- Missing start day and month, but year present: If the start of study medication occurred in the same year as the occurrence of the AE/concomitant medication, the start day and month will be assigned to the date of first intake of study medication. Otherwise the start day and month will be set to January 1st.
- Missing end day, but month and year present: The end day will be set to the last day of the month.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- Missing end day and month, but year present: If the maximum of the subject's date of study termination or the date equivalent to 70 days after the subject's last intake of study medication is the same year as the occurrence of the AE/Concomitant medication, then the end day and month will be set to the maximum of the date of study termination or the date equivalent to 70 days after last intake of study medication. Otherwise the end day will be set to December 31st of the given year.

Efficacy variables

The start and end dates of flares will be imputed for the analyses of AU flares data, in order to allow the calculation of time intervals between 2 consecutive flares and the duration of flares. Different rules will be applied for pre-study and on-study period flares.

For the interim analysis, the pre-study period is starting 48 weeks prior to Baseline, i.e. 336 days prior to Baseline until Baseline. The on-study period for the interim analysis is starting from Baseline until week 48, i.e. until study day 336.

For the final analysis, the pre-study period is starting **96 weeks** prior to Baseline, i.e. **672** days prior to Baseline until Baseline. The on-study period for the final analysis is starting from Baseline until week 96, i.e. until **study day 672**. In addition, **the period definition used for interim with period length of 48 weeks will be repeated for the final analysis**.

Pre-study period (**48 vs 96 weeks prior to Baseline, i.e. 336 vs 672 days prior to Baseline, until Baseline for interim vs final analysis):**

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Month and year are the same as month and year of the end date of previous flare, and the end date is clearly greater than end date of previous flare, then impute with the end date of previous flare.
  - In the case that month and year of the start date of pre-study period is the same month and year of the end date of previous flare, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Year is the same as year of the end date of previous flare, then impute with the end date of previous flare.
  - In the case that year of the start date of pre-study period is the same year of the end date of previous flare, then use the latest date of both for imputation.
- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred within the pre-study period. Then the maximum (latest) of start date of pre-study period and end date of previous flare will be considered.

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Month and year are the same as month and year of the start date of subsequent pre-study flare before Baseline, then impute with the start date of subsequent pre-study flare.
  - In the case that month and year of the end date of pre-study period is the same month and year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Year is the same as year of the start date of subsequent pre-study flare, then impute with the start date of subsequent pre-study flare.
  - In the case that year of the end date of pre-study period is the same year of the start date of subsequent pre-study flare, then use the earliest date of both for imputation.
- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred within the pre-study period. Then the minimum (earliest) of end date of pre-study period **and start date of subsequent pre-study flare will be considered**.

On-study period (Baseline until end of treatment):

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of study medication, then impute with the start date of study medication.
  - Month and year are the same as month and year of the end date of previous on-study flare, and the end date is clearly greater than end date of previous on-study flare, then impute with the end date of previous on-study flare.
  - In the case that month and year of the start date of study medication is the same month and year of the end date of previous on-study flare, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of study medication, then impute with the start date of study medication.
  - Year is the same as year of the end date of previous on-study flare, then impute with the end date of previous on-study flare.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

  - In the case that year of the start date of study medication is the same year of the end date of previous on-study flare, then use the latest date of both for imputation.

- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred after treatment start during the on-study period. Then the maximum (latest) of start date of study medication and end date of previous on-study flare eye will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of on-study period, then impute with the end date of on-study period.
  - Month and year are the same as month and year of the start date of subsequent on-study flare, then impute with the start date of subsequent on-study flare.
  - In the case that month and year of the end date of on-study period is the same month and year of the start date of subsequent on-study flare, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of on-study period, then impute with the end date of on-study period.
  - Year is the same as year of the start date of subsequent on-study flare, then impute with the start date of subsequent on-study flare.
  - In the case that year of the end date of on-study period is the same year of the start date of subsequent on-study flare, then use the earliest date of both for imputation.
- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred after treatment start during the on-study period. **Then the minimum (earliest) of end date of on-study period and start date of subsequent on-study flare will be considered.**

Has been changed to:

Analysis **is** based on observed case **only**.

If the Baseline value is missing, the Screening value may be used as the Baseline. In general, if no measurement is available prior to receiving study medication, then the Baseline value is treated as missing, and will not be imputed.

Premature withdrawal visits should be assigned to what would have been the next scheduled visit for the particular assessment that is being summarized. If a premature withdrawal visit occurs on the same date as a scheduled visit, then the assessments will be summarized with the scheduled visit.

Safety variables

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

For analyses of AEs and concomitant medication usage, a complete date must be established in order to correctly identify the AE or medication as occurring during treatment or not. For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the subject data listings (i.e., no imputed values will be displayed in data listings).

- Missing start day, but month and year present: If the start of study medication occurred in the same month and year as the occurrence of the AE/concomitant medication, the start day of the event/concomitant medication will be assigned to the day of first intake of study medication. Otherwise the start day will be set to the 1st day of the month.
- Missing start day and month, but year present: If the start of study medication occurred in the same year as the occurrence of the AE/concomitant medication, the start day and month will be assigned to the date of first intake of study medication. Otherwise the start day and month will be set to January 1st.
- If start day, month, and year are missing, start date will be replaced by the date of first study drug administration or by the first day of January of the year of the end date, if the first administration date is after the end date.
- Missing end day, but month and year present: The end day will be set to the last day of the month.
- Missing end day and month, but year present: If the maximum of the subject's date of study termination or the date equivalent to 70 days after the subject's last intake of study medication is the same year as the occurrence of the AE/Concomitant medication, then the end day and month will be set to the maximum of the date of study termination or the date equivalent to 70 days after last intake of study medication. Otherwise the end day will be set to December 31st of the given year.

Efficacy variables

The start and end dates of flares will be imputed for the analyses of AU flares data, in order to allow the calculation of time intervals between 2 consecutive flares and the duration of flares. Different rules will be applied for pre-study and on-study period flares.

For the interim analysis, the pre-study period is starting 48 weeks prior to Baseline, i.e. 336 days prior to Baseline until Baseline. The on-study period for the interim analysis is starting from Baseline until week 48, i.e. until study day 336.

For the final analysis, the pre-study period is starting **24 months** prior to Baseline, i.e. **730** days prior to Baseline until Baseline. The on-study period for the final analysis is starting from Baseline until week 96, i.e. until **the date of the Week 96 or discontinuation visit**. In addition, **sensitivity and supportive analyses use different definitions (as defined in Section 8.1.2).**

Pre-study period:

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- Month and year are the same as month and year of the start date of pre-study period, then impute with the start date of pre-study period.
- Month and year are the same as month and year of the end date of previous flare **in the same eye**, and the end date is clearly greater than end date of previous flare **in the same eye**, then impute with the end date of previous flare.
- In the case that month and year of the start date of pre-study period is the same month and year of the end date of previous flare **in the same eye**, then use the latest date of both for imputation.

- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of pre-study period, then impute with the start date of pre-study period.
  - Year is the same as year of the end date of previous flare **in the same eye**, then impute with the end date of previous flare.
  - In the case that year of the start date of pre-study period is the same year of the end date of previous flare **in the same eye**, then use the latest date of both for imputation.
- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred within the pre-study period. Then the maximum (latest) of start date of pre-study period and end date of previous flare **in the same eye** will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Month and year are the same as month and year of the start date of subsequent pre-study flare **in the same eye** before Baseline, then impute with the start date of subsequent pre-study flare.
  - In the case that month and year of the end date of pre-study period is the same month and year of the start date of subsequent pre-study flare **in the same eye**, then use the earliest date of both for imputation.
- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of pre-study period, then impute with the end date of pre-study period.
  - Year is the same as year of the start date of subsequent pre-study flare **in the same eye**, then impute with the start date of subsequent pre-study flare.
  - In the case that year of the end date of pre-study period is the same year of the start date of subsequent pre-study flare **in the same eye**, then use the earliest date of both for imputation.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred within the pre-study period **(of 2 years prior to baseline).** Then the minimum (earliest) of end date of pre-study period **(i.e. the day prior to treatment start date), the day prior to start date of subsequent pre-study flare in the same eye, or the start date plus the median duration of flare in that subject will be considered. The median duration of flare is determined from all flares recorded for that subject that have complete start and end dates (regardless of eye), including any flares prior to or during the pre-study period and on-study period.**

On-study period (Baseline until end of treatment):

- Missing start day, but month and year present: Impute the first day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the start date of study medication, then impute with the start date of study medication.
  - Month and year are the same as month and year of the end date of previous on-study flare **in the same eye**, and the end date is clearly greater than end date of previous on-study flare **in the same eye**, then impute with the end date of previous on-study flare **in the same eye**.
  - In the case that month and year of the start date of study medication is the same month and year of the end date of previous on-study flare **in the same eye**, then use the latest date of both for imputation.
- Missing start day and month, but year present: Impute January 1st unless one of the following conditions apply:
  - Year is the same as year of the start date of study medication, then impute with the start date of study medication.
  - Year is the same as year of the end date of previous on-study flare **in the same eye**, then impute with the end date of previous on-study flare **in the same eye**.
  - In the case that year of the start date of study medication is the same year of the end date of previous on-study flare **in the same eye**, then use the latest date of both for imputation.
- Completely missing start dates of flares will only be imputed if it is clear based on the end date that the flare occurred after treatment start during the on-study period. Then the maximum (latest) of start date of study medication and end date of previous on-study flare **in the same eye** will be considered.
- Missing end day, but month and year present: Impute the last day of the month unless one of the following conditions apply:
  - Month and year are the same as month and year of the end date of on-study period, then impute with the end date of on-study period.
  - Month and year are the same as month and year of the start date of subsequent on-study flare **in the same eye**, then impute with the start date of subsequent on-study flare **in the same eye**.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

- In the case that month and year of the end date of on-study period is the same month and year of the start date of subsequent on-study flare **in the same eye**, then use the earliest date of both for imputation.

- Missing end day and month, but year present: Impute December 31st unless one of the following conditions apply:
  - Year is the same as year of the end date of on-study period, then impute with the end date of on-study period.
  - Year is the same as year of the start date of subsequent on-study flare **in the same eye**, then impute with the start date of subsequent on-study flare **in the same eye**.
  - In the case that year of the end date of on-study period is the same year of the start date of subsequent on-study flare **in the same eye**, then use the earliest date of both for imputation.
- Completely missing end dates of flares will only be imputed if it is clear based on the start date that the flare occurred after treatment start during the on-study period. **Then the start date of subsequent on-study flare in the same eye will be considered. Otherwise it will be assumed to be ongoing.**

Changes in section 4.3

An additional interim data-cut might be needed to support a regulatory submission in nr-axSpA.

Has been changed to:

An additional interim data-cut might be needed to support a regulatory submission in nr-axSpA **or axSpA**.

Changes in section 4.5

Although there is no intention to stop the study early due to efficacy or futility on the basis of the interim results, **α=0.0001** will be spent in conjunction with this formal interim review of the data (Haybittle, 1971), and the final analysis of the primary efficacy variable will be conducted at the reduced 2-sided α-level of 0.049.

Has been changed to:

Although there is no intention to stop the study early due to efficacy or futility on the basis of the interim results, **α=0.001** will be spent in conjunction with this formal interim review of the data (Haybittle, 1971), and the final analysis of the primary efficacy variable will be conducted at the reduced 2-sided α-level of 0.049.

Changes in section 4.8

Not applicable

Has been changed to:

The primary efficacy variable will be summarized using the following subgroups:

- **Gender (Male, Female)**
- **Age (<median, >=median)**

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

- **axSpA subpopulation as determined by the investigator (Ankylosing Spondylitis, non-radiographic axial spondyloarthritis)**

A Poisson regression model as described for the primary analysis will be used including terms for the subgroup and the subgroup by treatment period interaction. The results will be presented, but no p-values will be given.

Changes in section 6.1

Tables with descriptive statistics and listings will be given for the demographic variables age (at time of informed consent), gender, weight and height at Baseline, race, ethnicity, body mass index (BMI), and body surface area (BSA). Demographic characteristics will be summarized for the SS.

The age will be calculated with the documented year of birth, January as imputed month and 1 as imputed day of birth, at the time of informed consent, as follows:

Has been changed to:

Tables with descriptive statistics and listings will be given for the demographic variables age (at time of informed consent), gender, weight and height at Baseline, race, ethnicity, body mass index (BMI), and body surface area (BSA) **as calculated below**. Demographic characteristics will be summarized for the SS.

Only the year of birth has been collected, so the age will be calculated with the documented year of birth, January as imputed month and 1 as imputed day of birth, at the time of informed consent, as follows:

Changes in section 6.2

Disease duration of axSpA will be calculated as the difference between the date of Baseline visit and the onset date of **axSpA** as documented on the medical history pages. If the onset date of axSpA is partial, it will be imputed to the most recent feasible date (ie, **last** day of the month if only day is missing, or the **last** day of the year if day and month are missing). Disease duration of axSpA will be summarized as continuous variable in **days**, as well as categorized into <2 years and ≥2 years.

Time since onset of first uveitis flare will be calculated as the difference between the date of Baseline visit and the onset date of first uveitis flare. If the onset date of first uveitis flare is partial, it will be imputed to the **last** feasible date (ie, first day of the month if only day is missing, or the first day of the year if day and month are missing). Time since onset of first uveitis flare will be summarized as continuous variable in **days**, as well as categorized into <2 years and ≥2 years.

Has been changed to:

Disease duration of axSpA will be calculated as the difference between the date of Baseline visit and the onset date of **Axial spondyloarthritis or Ankylosing spondylitis** as documented on the medical history pages. If the onset date of axSpA is partial, it will be imputed to the most recent feasible date (ie, **first** day of the month if only day is missing, or the **first** day of the year if day and month are missing). Disease duration of axSpA will be summarized as continuous variable in **years**, as well as categorized into <2 years and ≥2 years.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Time since onset of first uveitis flare will be calculated as the difference between the date of Baseline visit and the onset date of first uveitis flare. If the onset date of first uveitis flare is partial, it will be imputed to the **most recent** feasible date (ie, first day of the month if only day is missing, or the first day of the year if day and month are missing)**.** Time since onset of first uveitis flare will be summarized as continuous variable in **years**, as well as categorized into <2 years and ≥2 years.

If medical history terms of back pain or inflammatory back pain are missing for a subject, then the terms axial spondyloarthritis or Ankylosing spondylitis will used instead.

Changes in section 6.3

The Medical history contains any previous or ongoing medical conditions that occurred prior to study entry. They will be listed for all subjects in the **ES** including the reported term, start date (month and year only), end date (month and year only) (or ongoing if applicable) and the information whether the procedure is related to axSpA.

Has been changed to:

The Medical history contains any previous or ongoing medical conditions that occurred prior to study entry. They will be listed for all subjects in the **SS** including the reported term, start date (month and year only), end date (month and year only) (or ongoing if applicable) and the information whether the procedure is related to axSpA.

Changes in section 6.4

A medication can be defined as both prior and concomitant. Where a medication start or stop date is (partially) missing, the medication will be considered a concomitant medication, unless the partial data is suggesting otherwise. Imputation of dates will follow the rules described in section Section 4.2 . Concomitant medications will be summarized for subjects in the SS. The following information will be included in the listings: reported term, dose per intake and dose unit, frequency, formulation, route, indication, start and stop dates (including relative study day), and in case of missing stop date, the information whether the intake of medication is ongoing.

Intake of any new prohibited medication since last visit is collected and presented in a separate listing.

Has been changed to:

A medication can be defined as both prior and concomitant. Where a medication start or stop date is (partially) missing, the medication will be considered a concomitant medication, unless the partial data is suggesting otherwise. Imputation of dates will follow the rules described in section Section 4.2 . Concomitant medications will be summarized for subjects in the SS. The following information will be included in the listings: reported term, dose per intake and dose unit, frequency, formulation, route, indication, start and stop dates (including relative study day), and in case of missing stop date, the information whether the intake of medication is ongoing.

Past medication summaries will be generated for disease-modifying antirheumatic drugs (DMARDs), non-steroidal anti-inflammatory drugs (NSAIDs), and anti-tumor necrosis factors (anti-TNFs) by anatomical therapeutic chemical (ATC) Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT).

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Concomitant DMARDs and NSAIDs will be summarized by WHO-DD Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT). Concomitant medications other than DMARDs and NSAIDs will also be summarized.

Intake of any new prohibited medication since last visit is collected and presented in a separate listing.

Changes in section 7

There will be 2 approaches to calculate treatment compliance. The first will utilize the number of administered syringes and compare them to the scheduled expected number of syringes. Two syringes should be administered for the loading doses (at weeks 0, 2 and 4) however only 1 syringe should be administered at any other visit. The difference in number of syringes between the total actual used and the total expected syringes will be summarized. In addition, a ratio of compliance will be further computed based on the number of actual and expected syringes. The ratio of compliance will be summarized as a continuous variable and categorically (<0.80, ≥0.80-≤1.0 and >1.0). The general formula for the compliance ratio (CR) is given as follows:

CR = # actual syringes / # expected syringes

The second approach defines compliance with study drug administration based upon comparing the actual day of administration with the expected day of administration. The expected day of administration will be based upon the Baseline date. The sum of the absolute difference in days between the actual and expected days will be summarized. In addition, a ratio of compliance will be further computed based upon the actual and expected day. The ratio of compliance will also be summarized as a continuous variable and categorically (<0.80, ≥0.80-≤1.0 and >1.0). The general formula for the compliance ratio based on duration (CRD) is given as follows:

CRD = (Study Duration - Cumulative Difference) / Study Duration

The CRD ranges between 0 and 1. To calculate study duration, the date of the Week **96** visit or the last injection date prior to study treatment discontinuation will be compared to the Baseline date, as shown below:

Study Duration (days) = Week **96** visit/last injection date – Baseline **date (maximum value is 675 days)**

Cumulative Difference (days) = sum (ABS [actual date – scheduled date])

The sum will be calculated for the 11 visits from Week 0 (Baseline) to Week 84, and for all planned home administrations from week 6 to week 94. In case the actual day of administration is missing, the maximum deviation to the scheduled day will be assumed (ie, 14 days). If, for a scheduled day with 2 planned injections, the syringes were administered on 2 different days, the maximum day difference of the 2 actual dates to the scheduled date will be utilized.

Has been changed to:

There will be 2 approaches to calculate treatment compliance. The first will utilize the number of administered syringes and compare them to the scheduled expected number of syringes. Two syringes should be administered for the loading doses (at weeks 0, 2 and 4) however only 1 syringe should be administered at any other visit **from Week 6 until Week 94/last scheduled injection**. The difference in number of syringes between the total actual used and the total expected syringes will be summarized. In addition, a ratio of compliance will be further

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

computed based on the number of actual and expected syringes. The ratio of compliance will be summarized as a continuous variable and categorically (<0.80, ≥0.80-≤1.0 and >1.0). The general formula for the compliance ratio (CR) is given as follows:

$$CR = \frac{N_{Syr_act}}{N_{Syr_exp}}$$

where N_{Syr_act} is the number of actual injections and N_{Syr_exp} is the number of expected injections. The second approach defines compliance with study drug administration based upon comparing the actual day of administration with the expected day of administration. The expected day of administration will be based upon the Baseline date **and the previous injection date**. The sum of the absolute difference in days between the actual and expected days will be summarized. In addition, a ratio of compliance will be further computed based upon the actual and expected day. The ratio of compliance will also be summarized as a continuous variable and categorically (<0.80, ≥0.80-≤1.0 and >1.0). The general formula for the compliance ratio based on duration (CRD) is given as follows:

$$CRD = \frac{T_{Dur_stu} - \Delta_{Cum}}{T_{Dur_stu}}$$

where T_{Dur_stu} is the study duration and Δ_{Cum} is the cumulative difference between actual and scheduled day of administration.

Specific details of this formula are provided in Appendix 12.1 (Compliance Ratio Calculation).

The CRD ranges between 0 and 1. To calculate study duration, the date of the Week **94** visit or the last injection date prior to study treatment discontinuation will be compared to the Baseline date, as shown below:

Study Duration (days) = Week 94 visit/last injection date – Baseline date + 14

The sum will be calculated for the 11 visits from Week 0 (Baseline) to Week 84, and for all planned home administrations from week 6 to week 94. In case the actual day of administration is missing, the maximum deviation to the scheduled day will be assumed (ie, 14 days). If, for a scheduled day with 2 planned injections, the syringes were administered on 2 different days, the maximum day difference of the 2 actual dates to the scheduled date will be utilized.

Changes in section 8

The efficacy analysis will be based on all subjects in the SS or FAS, as appropriate. If there is no missing AU flare incidence data from pre-study period, the FAS will be the same as the SS. **If there is missing AU flare incidence data from the pre-study period, the FAS will be used instead of the SS.** The primary efficacy analysis will additionally be repeated for the PPS.

Has been changed to:

The efficacy analysis will be based on all subjects in the SS or FAS, as appropriate. If there is no missing AU flare incidence data from pre-study period, the FAS will be the same as the SS. **The FAS will be used for analyses of the primary variable and the SS will be used for all other efficacy analyses.** The primary efficacy analysis will additionally be repeated for the PPS.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Changes in section 8.1

The primary efficacy analysis will consist of a comparison of the frequency of AU flares during the pre-study/historical period with that observed while in the study during CZP treatment. It is assumed that the frequency of AU flares follows a Poisson distribution. As such, the analysis will be performed as a generalized estimating equations analysis for Poisson outcome that will take into account the possible within-subject correlation (between the retrospective and prospective AU flare counts). The model will contain 2 records per subject corresponding to the frequency of AU flares before study and during the study, offset by log-time of each of the reporting periods. Disease duration of axSpA will be included as a factor in the model, initially defined as <2 years and ≥2 years; the specific cutoffs for analysis may deviate from this based on the distribution seen in the data. The p-value for effect of CZP treatment on frequency of AU flares, along with rate ratio (CZP/historical) and 95% confidence interval, will be obtained from this model.

The primary efficacy analysis hypothesis will be tested based on significance level of alpha 0.05 and will be 2-sided. The analysis of the other secondary efficacy parameters will be supportive. Therefore, no alpha adjustment for secondary efficacy variables will be performed. The alpha adjustment related to the formal interim review of data is described in Section 4.5.

Statistical Model

Let Y_{ij} be the number of AU flare episodes for subject i in period j, where t_{ij} is the length of the interval. It is assumed that the responses within a subject are correlated. For the statistical model assumptions about the correlation structure need to be done. Assuming an exchangeable correlation structure means that correlation between any two responses of the ith individual is the same. Also we account for different interval length by adding the offset term $\log(t_{ij})$. The mean of the Poisson distribution is $E(Y_{ij})=\mu_{ij}$, and the mean response per unit time is μ_{ij}/t_{ij}. The Poisson model using a log-link can be written as

$$\mathbf{Log(\mu_{ij}) = \beta_0 + \beta_1 * x1_{ij} + \beta_2 * x2_{ij} + log(t_{ij})}$$

which resolves to $\mathbf{\mu_{ij} = t_{ij} * exp(\beta_0) * exp(\beta_1 * x1_{ij}) * exp(\beta_2 * x2_{ij})}$**,** with the variables

$x1_{ij}$: variable for period: 1 on-study, 0 pre-study

$\mathbf{x2_{ij}}$**:** variable for disease duration of axSpA: <2 years, ≥2 years

Has been changed to:

The primary efficacy analysis will consist of a comparison of the frequency of AU flares during the pre-study/historical period with that observed while in the study during CZP treatment. It is assumed that the frequency of AU flares follows a Poisson distribution. As such, the analysis will be performed as a generalized estimating equations analysis for Poisson outcome that will take into account the possible within-subject correlation (between the retrospective and prospective AU flare counts). The model will contain 2 records per subject corresponding to the frequency of AU flares before study and during the study, offset by log-time of each of the reporting periods. Disease duration of axSpA will be included as a factor in the model, initially defined as <2 years and ≥2 years; the specific cutoffs for analysis may deviate from this based on the distribution seen in the data. The p-value for effect of CZP treatment on frequency of AU flares, along with rate ratio (CZP/historical) and 95% confidence interval, will be obtained from

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

this model. **The rate per 96 weeks and rate per year for each study period from the model will also be presented.**

The primary efficacy analysis hypothesis will be tested based on significance level of alpha 0.05 and will be 2-sided. The analysis of the other secondary efficacy parameters will be supportive. Therefore, no alpha adjustment for secondary efficacy variables will be performed. The alpha adjustment related to the formal interim review of data is described in Section 4.5.

Statistical Model

Let Y_{ij} be the number of AU flare episodes for subject i in period j, where t_{ij} is the length of the interval. It is assumed that the responses within a subject are correlated. For the statistical model assumptions about the correlation structure need to be done. Assuming an exchangeable correlation structure means that correlation between any two responses of the ith individual is the same. Also we account for different interval length by adding the offset term $\log(t_{ij})$. The mean of the Poisson distribution is $E(Y_{ij})= \mu_{ij}$, and the mean response per unit time is μ_{ij}/t_{ij}. The Poisson model using a log-link can be written as

$$\log \mu_{ij} = \beta_0 + \beta_1 x_{1;ij} + \beta_2 x_{2;i} + \log t_{ij}$$

which resolves to

$$\mu_{ij} = t_{ij} \exp(\beta_0) * \exp(\beta_1 x_{1;ij}) * \exp(\beta_2 x_{2;i})$$

with the variables

$x1_{ij}$: variable for period: 1 on-study, 0 pre-study

$x2_i$: variable for disease duration of axSpA: <2 years, ≥2 years

Changes in section 8.1.1

Pre-study and on-study AU flare episodes will be counted. An episode of AU flares is considered as pre-study, if the **start date of flare** is before the start date of study medication and **if the flare occurs within the 96 weeks (672 days)** prior to the start of study medication. For the interim analysis, an episode of AU flares is considered as pre-study if the flare occurs within 48 weeks (336 days) prior to the start of study medication. An episode of AU flares is considered on-study, if the start date of flare is on or after the start date of study medication. Pre-study and on-study flares of the same eye will be combined if the time interval between the end of pre-study flare and the start of subsequent on-study flare is less than 3 months (90 days). The resulting flare episode will be considered as pre-study flare episode.

The log-time offset in the Poisson model will consist of the length of pre-study and on-study reporting periods in days. As the documentation of pre-study episodes is extended to up to 2 years prior to Baseline, the pre-study period length for offset is **672 days for each subject.** The on-study period length for offset is the time in days from start of study medication until last visit. For the interim analysis, an offset of 336 days will be used for the pre-study period, and the time

REDACTED COPY
This document cannot be used to support any marketing authorization applications and any extensions or variations thereof.

in days from start of study medication until week 48 visit or early termination visit where applicable will be used.

Disease duration of axSpA, which will be included in the Poisson model as a factor, is calculated **from the start date of first AU flare episode until the Baseline visit date**, and is categorized into <2 years and ≥2 years.

Has been changed to:

Pre-study and on-study AU flare episodes will be counted. An episode of AU flares is considered as pre-study, if **the start date of flare** is before the start date of study medication and **if the start date of flare** occurs within the **24 months (730 days)** prior to the start of study medication. For the interim analysis, an episode of AU flares is considered as pre-study if the flare occurs within 48 weeks (336 days) prior to the start of study medication**.**

For the Week 96 (primary) analysis, an episode of AU flares is considered on-study, if the start date of flare is on or after the start date of study medication **and the start date of the flare occurs up to and including the last visit in the treatment period, i.e. Week 96 or discontinuation visit.**

Pre-study and on-study flares of the same eye will be combined if the time interval between the end of pre-study flare and the start of subsequent on-study flare is less than 3 months (90 days). The resulting flare episode will be considered as pre-study flare episode.

The log-time offset in the Poisson model will consist of the length of pre-study and on-study reporting periods in days. As the documentation of pre-study episodes is extended to up to 2 years prior to Baseline, the pre-study period length for offset is **730 days for each subject, unless the date of first uveitis flare is within 2 years of baseline, in which case the pre-study period length will be from date of first uveitis flare until baseline.** The on-study period length for offset is the time in days from start of study medication until last visit**, i.e. Week 96 or discontinuation visit.** For the interim analysis, an offset of 336 days will be used for the pre-study period **or time since first uveitis flare if shorter than 336 days,** and the time in days from start of study medication until week 48 visit **(taken as 336 days**) or early termination visit **(taken as 14 days after the last treatment)** where applicable will be used **for the on-study period.**

Disease duration of axSpA, which will be included in the Poisson model as a factor, is calculated **as defined in Section 6.2,** and is categorized into <2 years and ≥2 years.

Changes in section 8.1.2

The AU flare rate (event rate) per 100 patient-years of exposure, both before and during the study will be calculated. The corresponding rate ratio along with 95% confidence interval (Garwood, 1936) will be presented as a key supportive analysis to the primary efficacy analysis results (see Section 8.2.1).

Has been changed to:

The primary analysis will be repeated using the pre-study period as 96 weeks (672 days) prior to start of study medication for counting flares and as the offset, instead of 24 months. The on-study period will be the same as that used for the primary analysis.

This document cannot be used to support any marketing authorisation application and any extensions or variations thereof.
REDACTED COPY

In addition, the analysis for the full 2 years before and after study medication will be provided. The pre-study period will be the 2 years (730 days) prior to start of study medication, which will be used for counting flares and as the offset. The on-study period will be determined for whose episodes of AU flares for which the start date of flare is on or after the start date of study medication and up to and including the SFU visit or date of last administration of study medication + 70 days.

An analysis for the first 48 weeks will also be produced. The period for pre-study will be the 12 months (365 days) prior to start of study medication, which will be used for counting flares and as the offset. An episode of AU flares will be considered on-study, if the start date of flare is on or after the start date of study medication and up to and including the Week 48 visit or discontinuation visit if prior to Week 48. The exact date of the Week 48 or discontinuation visit will be used for each subject. This analysis will also be repeated using 48 weeks (336 days) as the pre-study period for counting flares and as offset, but keeping the on-study period of 48 weeks the same (i.e. using the exact date of the Week 48 or discontinuation visit).

The primary efficacy analysis will be repeated in the subgroup of subjects excluding those with any pre-study flares with completely missing end dates in the 2 years prior to baseline. Subjects will be excluded if any of the flares that started within the 2 years prior to baseline have completely missing end dates (i.e. no day, month or year).

Changes in section 8.2.1

Number of AU flares

Has been changed to:

Number of AU flares per 100 patient-years

Changes in section 8.2.2

[redacted], PtGADA, [redacted] are all assessed on a numerical scale (0 to 10 units) (Lukas et al, 2009).

The sum of these weighted components gives the ASDAS.

If 1 component for the ASDAS is missing at a given visit, that component will be imputed by carrying the last observation forward, and the ASDAS will be calculated accordingly. If more than one component for the ASDAS is missing, ASDAS will be treated as missing.

For CRP values below the lower limit of quantification (<4mg/L), half the lower limit (2mg/L) will be used as the imputed value.

The nomenclature of disease activity was updated by ASAS (Machado et al, 2018). The state 'Moderate Disease Activity' was replaced by 'Low Disease Activity' state. Disease activity categories based on ASDAS are as follows:

- ASDAS-Inactive Disease (ASDAS-ID): ASDAS <1.3
- ASDAS-Low Disease Activity (ASDAS-LD): ASDAS >=1.3, <2.1
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS >=2.1, <=3.5

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extension or variations thereof.

- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

The variables related to ASDAS clinical improvement are defined as follows:

- **ASDAS-Clinical Improvement (ASDAS-CI): ASDAS reduction (improvement) of >0 relative to Baseline**
- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of >=1.1 relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of >=2.0 relative to Baseline **or has the lowest score possible post-baseline (i.e. when CRP<LLOQ and all other components are 0, then the minimum ASDAS score is 0.636 to 3 decimal places)**.

Has been changed to:

[redacted], PtGADA, [redacted] **are all assessed on a numerical scale (0 to 10 units) (Lukas et al, 2009).**

The sum of these weighted components gives the ASDAS.

If 1 component for the ASDAS is missing at a given visit, that component will be imputed by carrying the last observation forward, and the ASDAS will be calculated accordingly. If more than one component for the ASDAS is missing, ASDAS will be treated as missing.

For CRP values below the lower limit of quantification (<4mg/L), half the lower limit (2mg/L) will be used as the imputed value. **This approach will be used for summaries of CRP as well.**

The nomenclature of disease activity was updated by ASAS (Machado et al, 2018). The state 'Moderate Disease Activity' was replaced by 'Low Disease Activity' state. Disease activity categories based on ASDAS are as follows:

- ASDAS-Inactive Disease (ASDAS-ID): ASDAS <1.3
- ASDAS-Low Disease Activity (ASDAS-LD): ASDAS >=1.3, <2.1
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS >=2.1, <=3.5
- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

The variables related to ASDAS clinical improvement are defined as follows:

- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of >=1.1 relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of >=2.0 relative to Baseline.

Changes in sections 8.2.5, 8.2.6, 8.2.7, 8.2.8

The data including the change from baseline at Week 48 and Week 96 will be summarized.

Has been changed to:

The data including the change from baseline at Week 48 and Week 96 will be summarized **as secondary endpoints**.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Changes in section 8.3

Has been addeed:

All other efficacy endpoints with exception of flare duration and severity will be tabulated at all time points acc to assessment schedule.

Changes in section 8.3.1

For each subject in FAS, the **mean** duration **of on-study AU flares,** and the maximum severity of on-study AU flares will be determined and summarized**.** The duration of an episode of AU flare is calculated as follows:

Duration (days) = end date of flare – start date of flare +1.

For uveitis flare, the highest grade (severity) noted during an episode of uveitis will be graded as follows:

Grade 0=None; Grade 1+=Faint; Grade 2+=Moderate, iris and lens details clear; Grade3+=Marked, iris and lens details hazy; Grade4+=Intense, fibrin or plastic aqueous; ND=Not done or not assessed; UNK=Unknown

Has been changed to:

For each subject in FAS, the **total** duration and the maximum severity of **pre-study and** on-study AU flares will be determined and summarized. The duration of an episode of AU flare **(T_{Dur_Fl})** is calculated as follows:

$$T_{Dur_Fl} = Date_{End_Fl} - Date_{Start_{Fl}} + 1$$

with $Date_{End_Fl}$ and $Date_{Start_FL}$ the end date and start date of flare, respectively.

The duration of pre-study and on-study periods is defined in Section 4.2.

For the interim analysis, the pre-study period is starting 48 weeks prior to Baseline, i.e. 336 days prior to Baseline until Baseline. The on-study period for the interim analysis is starting from Baseline until week 48, i.e. until study day 336.

For the final analysis, the pre-study period is starting 24 months prior to Baseline, i.e. 730 days prior to Baseline until Baseline. The on-study period for the final analysis is starting from Baseline until week 96, i.e. until the date of the Week 96 or discontinuation visit.

The highest grade (severity) of uveitis flare across all flares per subject will be summarized. The highest grade (severity) noted during an episode of uveitis will be graded as follows:

Grade 0=None; Grade 1+=Faint; Grade 2+=Moderate, iris and lens details clear; Grade3+=Marked, iris and lens details hazy; Grade4+=Intense, fibrin or plastic aqueous; ND=Not done or not assessed; UNK=Unknown.

Changes in section 8.3.2

Extra-articular assessments (eg, number of IBD exacerbations **[in subjects with concurrent IBD**] and number of psoriasis exacerbations [**in subjects with concurrent psoriasis**]) are to be performed at Screening, Baseline, Weeks 12, 24, 36, 48, 60, 72, 84, and at Completion at Week 96/WD Visit. At each visit, the number of episodes since the last visit and the number of days

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

with IBD/psoriasis are collected. Episodes are considered equivalent to exacerbations. These data will be summarized using summary statistics for continuous variables.

Has been changed to:

Extra-articular assessments (eg, number of IBD exacerbations and number of psoriasis exacerbations) are to be performed at Screening, Baseline, Weeks 12, 24, 36, 48, 60, 72, 84, and at Completion at Week 96/WD Visit. At each visit, the number of episodes since the last visit and the number of days with IBD/psoriasis are collected. Episodes are considered equivalent to exacerbations. These data will be summarized using summary statistics for continuous variables.

Changes in section 8.3.5:

For the calculation of the SF-36 domain scores and the component summaries PCS and MCS, the scoring software QualityMetric Health Outcomes™ Scoring Software 4.5 will be used. The norm-based scores (based on the US general population) will be utilized for analysis.

Has been changed to:

For the calculation of the SF-36 domain scores and the component summaries PCS and MCS, the scoring software QualityMetric Health Outcomes™ Scoring Software 4.5 will be used. The norm-based scores (based on the US general population **for 2009**) will be utilized for analysis.

Changes in section 10.1:

The number of doses received will be summarized with continuous statistics for the SS.

The duration of exposure is calculated as follows:

Exposure duration = Date of last dose of study medication – date of first dose of study medication + 14 days

The cumulative duration of exposure will be summarized for subjects exposed for at least one day, ≥183 days, ≥365 days, ≥548 days, ≥730 days.

Additionally, the subject time at risk will be summarized. Subject time at risk (in days) is calculated as:

Date of last administration of study medication – date of first administration + 70 days (except where a change of treatment or last clinical contact occurs prior to completion of this 70-day period. Date of last clinical contact per subject is the defined as the maximum of [last visit date including safety follow-up visits, last imputed AE start date, last date of termination or completion, last date of study drug administration, clinical cutoff date for ongoing studies, death date])**.**

In cases where the date of last clinical contact occurs prior to completion of this **70-day period**, subject time at risk (in days) will be calculated as:

Date of last clinical contact – date of first administration+1.

The study medication duration will be listed together with the dates of first and last study medication administration.

Has been changed to:

The number of doses received will be summarized with continuous statistics for the SS.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Duration of exposure to study medication will be calculated as:

$$Dur_{Exp} = date_{Last} - date_{First} + d_{\text{int}}$$

where d_{int} represents the dosing interval of 14 days.

The cumulative duration of exposure will be summarized for subjects exposed for at least one day, ≥183 days, ≥365 days, ≥548 days, ≥730 days.

Additionally, the subject time at risk will be summarized. Subject time at risk (in days) is calculated as:

$$T_{Risk} = date_{Last} - date_{First} + \left(5 * t_{1/2}\right)$$

where $t_{1/2}$ is 14 days.

If there is a change of treatment or the last clinical contact occurs prior to completion of the follow-up period, the date of last clinical contact per subject is defined as the maximum of [last visit date including safety follow-up visits, last imputed AE start date, last date of termination or completion, last date of study drug administration, death date] or the clinical cutoff date for ongoing studies if earlier than this maximum date).

In cases where the date of last clinical contact (**LastCC**) occurs prior to completion of this **follow-up period**, subject time at risk (in days) will be calculated as:

$$T_{Risk} = date_{LastCC} - date_{First} + 1$$

Subject years at risk is the sum of all relevant subject time at risk (in days) divided by 365.25.

The study medication duration will be listed together with the dates of first and last study medication administration.

Changes in section 10.2:

Tables showing TEAEs, serious TEAEs, non-serious TEAEs, TEAEs leading to discontinuation of study drug, TEAEs by maximum intensity, TEAEs by relationship, serious TEAEs by relationship, fatal TEAEs by relationship, TEAEs including subject numbers will be provided on the SS.

Additionally, more detailed summaries of TEAEs will be presented, which include the exposure adjusted incidence rate (EAIR) with associated 95% confidence interval, and the exposure adjusted event rate (EAER). **For EAIR, the numerator will be the total number of subjects experiencing the AE of interest. The denominator will be 100 patient-years; that is, the total summation of individual patient-years at risk up to the first occurrence of the AE of interest for subjects with that AE, and the total patient-years at risk for those subjects not experiencing that AE, divided by 100. EAIRs will be presented with a 95% exact confidence interval based upon the Chi-Square distribution (Ulm, 1990). For EAERs, the numerator will be the number of AEs including repeat occurrences in individual subjects; the denominator will be 100 patient-years. That is, the total summation of individual patient-years at risk divided by 100. No confidence interval will be computed for EAER.**

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Subject time at risk represents the time a subject was at risk for having an AE. The derivation is described in Section 10.1 .Subject years at risk is the sum of all relevant subject time at risk (in days) divided by 365.25.

Where an AE start date is (partially) missing, the AE will be considered TE if possible. Imputation rules provided in Section 4.2 are based on the assumption that if the start date is incomplete, either day only is missing, or day and months are missing, or date is completely missing. Although the algorithms for treatment-emergence depend on the onset date, imputation rules are provided for resolution date as well, as these may be needed for certain statistical analyses, such as an analyses of AE prevalence or AE duration.

Listings based on SS for all AEs will be provided. A glossary for the reported terms will also be generated utilizing SOC, HLT, and PT.

Anticipated SAEs, as described in Section 10.5 of the clinical study protocol, will not be summarized either by any additional tables or listings, as they are included in the standard AE summaries.

AEs of interest will be summarized. The following AEs of interest will be summarized in stand-alone tables:

- Opportunistic infections
- Malignant **and** unspecified tumors
- Malignant tumors
- Serious cardiovascular events
- Demyelinating-like disorders
- Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia
- Serious bleeding events

For the following AEs of interest, no additional tables will be provided, as these are included in the standard AE summaries:

- Serious infections
- Congestive heart failure
- Lupus and lupus-like illness
- Serious skin reactions (eg, Stevens Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme)

The following notes to AEs of interest in the clinical study protocol are given for reporting purposes only, and therefore will not be summarized separately: A confirmed LTBI must be reported as an AE of interest. Confirmed active TB is an AE of interest. In the event that an active TB occurrence meets the protocol's SAE criteria, it is also an SAE. Potential Hy's Law, defined as ≥3xULN alanine aminotransferase or aspartate aminotransferase with coexisting ≥2xULN total bilirubin in the absence of ≥2xULN alkaline phosphatase, with no alternative explanation for the biochemical abnormality is also considered an AE of interest.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

The following events are not officially considered to be AEs of interest but are nonetheless considered to be interesting enough to be summarized in stand-alone tables: 1) Hepatic events and 2) Hypersensitivity reactions and anaphylactic reactions.

The approach for summarizing AEs of interest is detailed in the UCB AE of interest program specifications for Cimzia.

Data collected on hospitalization will be presented in a listing.

Has been changed to:

Tables showing TEAEs, serious TEAEs, non-serious TEAEs, TEAEs leading to discontinuation of study drug, TEAEs by maximum intensity, TEAEs by relationship, serious TEAEs by relationship, fatal TEAEs by relationship, TEAEs including subject numbers will be provided on the SS.

Additionally, more detailed summaries of TEAEs will be presented, which include the exposure adjusted incidence rate (EAIR) with associated 95% confidence interval, and the exposure adjusted event rate (EAER).

The EAIR is defined as the number of subjects (n) with a specific AE adjusted for the exposure and will be scaled to 100 patient-years:

$$EAIR = 100 \cdot n \Big/ \sum_{i=1}^{N} T_{Exp,i}$$

If a subject has multiple events, the time of exposure is calculated to the first occurrence of the AE of interest. If a subject has no events, the total time at risk is used.

Exact Poisson 95% confidence intervals for incidence rates are calculated using the relationship between the Poisson and the Chi-square distribution (Ulm, 1990):

$$LCL = \chi^2_{2n,\alpha/2} \Big/ 2$$

$$UCL = \chi^2_{2(n+1),1-\alpha/2} \Big/ 2$$

where n is the number of subjects with a specific AE for the incidence rate of interest and is the basis for the number of the degrees of freedom for the chi-square quantile for the upper tail probability χ^2.

The EAER will be the number of AEs including repeat occurrences in individual subjects divided by the total time at risk scaled to 100 patient-years and calculated using:

$$EAER = 100 \cdot N_{AE} \Big/ \sum_{i=1}^{N} T_{Risk,i}$$

where N_{AE} is the total number of AEs.

No confidence interval will be computed for EAER.

Where an AE start date is (partially) missing, the AE will be considered TE if possible. Imputation rules provided in Section 4.2 are based on the assumption that if the start date is incomplete, either day only is missing, or day and months are missing, or date is completely missing. Although the algorithms for treatment-emergence depend on the onset date, imputation

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

rules are provided for resolution date as well, as these may be needed for certain statistical analyses, such as an analyses of AE prevalence or AE duration.

Listings based on SS for all AEs will be provided. A glossary for the reported terms will also be generated utilizing SOC, HLT, and PT.

Anticipated SAEs, as described in Section 10.5 of the clinical study protocol, will not be summarized either by any additional tables or listings, as they are included in the standard AE summaries.

AEs of interest will be summarized. The following AEs of interest will be summarized in stand-alone tables:

- Opportunistic infections
- Malignant **or** unspecified tumors
- Malignant tumors
- Serious cardiovascular events
- Demyelinating-like disorders
- Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia
- Serious bleeding events

For the following AEs of interest, no additional tables will be provided, as these are included in the standard AE summaries:

- Serious infections
- Congestive heart failure
- Lupus and lupus-like illness
- Serious skin reactions (eg, Stevens Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme)

The following notes to AEs of interest in the clinical study protocol are given for reporting purposes only, and therefore will not be summarized separately: A confirmed LTBI must be reported as an AE of interest. Confirmed active TB is an AE of interest. In the event that an active TB occurrence meets the protocol's SAE criteria, it is also an SAE. Potential Hy's Law, defined as ≥3xULN alanine aminotransferase or aspartate aminotransferase with coexisting ≥2xULN total bilirubin in the absence of ≥2xULN alkaline phosphatase, with no alternative explanation for the biochemical abnormality is also considered an AE of interest.

The following events are not officially considered to be AEs of interest but are nonetheless considered to be interesting enough to be summarized in stand-alone tables: 1) Hepatic events and 2) Hypersensitivity reactions and anaphylactic reactions.

The approach for summarizing AEs of interest is detailed in the UCB AE of interest program specifications for Cimzia **from 18 June 2018**.

Data collected on hospitalization will be presented in a listing.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Changes in section 10.3:

The changes from Baseline in laboratory evaluations will be analyzed over time in the SS for observed cases. Shift tables concerning the normal range at each post-baseline visit, will also be produced for each hematology and biochemistry laboratory **parameter**. The shifts will be categorized using L, N, H, missing, and total.

If a repeat sample is taken, the repeated sample data will be used when possible. If the value of a **parameter** collected at the scheduled visit is missing and an additional sample associated with this visit is taken before the next scheduled visit, the missing value will be replaced by this value. Early Withdrawal Visits will be assigned to what would have been the next scheduled visit. In the Urinalysis, microscopy (WBC, RBC, casts, crystals, and bacteria) will be performed only when there are abnormalities on the dipstick. This data will be listed and not presented in tables.

Has been changed to:

The changes from Baseline in laboratory evaluations will be analyzed over time in the SS for observed cases. Shift tables concerning the normal range at each post-baseline visit, will also be produced for each hematology and biochemistry laboratory **variable.** The shifts will be categorized using L, N, H, missing, and total.

If a repeat sample is taken, the repeated sample data will be used when possible. If the value of a **variable** collected at the scheduled visit is missing and an additional sample associated with this visit is taken before the next scheduled visit, the missing value will be replaced by this value. Early Withdrawal Visits will be assigned to what would have been the next scheduled visit. In the Urinalysis, microscopy (WBC, RBC, casts, crystals, and bacteria) will be performed only when there are abnormalities on the dipstick. This data will be listed and not presented in tables.

Changes in section 12.1:

Not applicable.

Has been changed to:

12.1 APPENDIX 1: Compliance Ratio Calculation

The general formula for treatment compliance is given as follows:

$$CR = \frac{T_{Dur_stu} - \Delta_{Cum}}{T_{Dur_stu}}$$

where T_{Dur_stu} is the study duration and Δ_{Cum} is the cumulative difference between actual and scheduled day of administration. The CRD ranges between 0 and 1.

There are 2 methods to determine the CR. CR(1) is based upon scheduled days which reference the Baseline visits. CR(2) is based upon scheduled days which reference the previous dosing administration date. CR(1) and CR(2) each will be computed separately. The definition of study duration and cumulative difference varies for the 2 calculations. In all cases, if the subject was randomized, but study drug was never administered then the CR is equal to 0. The CR is limited

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

to a range of 0 to 1. Therefore, if for any reason a negative value is computed, the CR must be set to 0. Specific details for each calculation follow.

CR(1):

$$T_{Dur_stu} = Date_{Last} - Date_{BL}$$
$$\Delta_{Cum} = \sum_i |AD_i - SD_i|$$

where i=Week 0, 2, 4, 6, … , XX

where AD_i = Actual day of administration, for Week i=0 to XX by 2
SD_i = (i/2) x 14, for Week i=0 to XX by 2

Note: XX represents the last completed visit in which study drug should have been administered. If subject is still in the study but AD_i is missing, then ABS[AD_i – SD_i]=14 for the missed Visit i.

CR(2):

$$T_{Dur_stu} = Date_{Last} - Date_{BL}$$

$$\Delta_{Cum} = \sum_i |AD_i - SD_i|$$

where i= Week 0, 2, 4, 6, … , XX

where AD_i = Actual day of administration, for Week i=0 to XX by 2
SD_i = {0, for Week i=0 }
{14 + AD_{i-2}, for Week i=2 to XX by 2 }

Note: XX represents the last completed visit in which study drug should have been administered. If subject is still in the study but AD_i is missing, then ABS(AD_i – SD_i)=14 for the missed Visit i.

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.
REDACTED COPY

Example

Week	Scheduled day (1)	Scheduled day (2)	Actual day	ABS Diff (1)	ABS Diff (2)
0	0	0	0	0	0
2	14	14	13	1	1
4	28	27	29	1	2
6	42	43	-	14	14
8	56	57	55	1	2
10	70	69	69	1	0
Total	-	-	-	18	19
CR				66/84 =0.79	65/84 =0.77

Note that this example is for a 12 week study.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

This document has been reviewed and approved per the Review and Approval of Clinical Documents Standard Operating Procedures. Signatures indicate that the final version of the Statistical Analysis Plan (SAP) or amended SAP is released for execution.

REDACTED COPY
This document cannot be used to support any marketing authorization application and any extensions or variations thereof.