

## STATISTICAL ANALYSIS PLAN

**Study: AS0006 (C-AXSPAND)**

**Product: Certolizumab pegol**

PHASE 3, MULTICENTER, RANDOMIZED, PLACEBO-CONTROLLED, DOUBLE-BLIND STUDY TO EVALUATE EFFICACY AND SAFETY OF CERTOLIZUMAB PEGOL IN SUBJECTS WITH ACTIVE AXIAL SPONDYLOARTHRITIS (AXSPA) WITHOUT X-RAY EVIDENCE OF ANKYLOSING SPONDYLITIS (AS) AND OBJECTIVE SIGNS OF INFLAMMATION

<b>SAP/Amendment Number</b>	<b>Date</b>
Final SAP	14 Jan 2015
SAP Amendment 1	22 Apr 2016
SAP Amendment 2	08 June 2017
SAP Amendment 3	26 March 2018
SAP Amendment 4-Draft	29 Jan 2019
SAP Amendment 4-Updated	22 Feb 2019
SAP Amendment 4	11 Mar 2019
SAP Amendment 5-Draft	20 Mar 2020
SAP Amendment 5	02 Apr 2020
SAP Amendment 6-Draft	03 Jun 2020
SAP Amendment 6	10 Jun 2020

### 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.....	6
1 INTRODUCTION .....	12
2 PROTOCOL SUMMARY .....	12
2.1 Study objectives .....	12
2.1.1 Primary objective.....	12
2.1.2 Secondary objectives .....	12
2.1.3 Other objectives .....	12
2.2 Study variables.....	13
2.2.1 Efficacy variables .....	13
2.2.1.1 Primary efficacy variable .....	13
2.2.1.2 Secondary efficacy variables.....	13
2.2.1.3 Other efficacy variables.....	14
2.2.2 Pharmacokinetic, exploratory biomarker, and pharmacogenomics variables .....	16
2.2.2.1 Primary Pharmacokinetic variables.....	16
2.2.2.2 Other biomarker variables .....	16
2.2.2.3 Other pharmacogenomics variables .....	16
2.2.3 Other immunological variables.....	16
2.2.4 Safety variables.....	17
2.2.4.1 Secondary safety variables .....	17
2.2.4.2 Other safety variables .....	17
2.3 Study design and conduct .....	23
2.3.1 Study periods .....	23
2.3.2 Study duration per subject .....	24
2.3.3 End of periods and study .....	24
2.3.4 Planned number of subjects and sites .....	25
2.3.5 Anticipated regions and countries .....	25
2.3.6 Stratification .....	25
2.4 Determination of sample size.....	25
3 DATA ANALYSIS CONSIDERATIONS .....	25
3.1 General presentation of summaries and analyses .....	25
3.2 General study level definitions .....	26
3.2.1 Analysis time points .....	26
3.2.1.1 Pre-treatment period .....	26
3.2.1.2 Double-blind treatment period .....	27
3.2.1.3 Open-label period.....	27
3.2.1.4 Follow-up period .....	28
3.2.1.5 Pre-SFE period .....	28

---

3.2.1.6	SFE period	28
3.2.2	Relative day	29
3.3	Definition of Baseline values	29
3.4	Protocol deviations	30
3.5	Analysis sets	30
3.5.1	Enrolled Set	30
3.5.2	Randomized Set	30
3.5.3	Safety Set	30
3.5.4	Full Analysis Set	30
3.5.5	Per Protocol Set	30
3.5.6	SFE Safety Set	30
3.5.7	SFE Full Analysis Set	31
3.6	Treatment assignment and treatment groups	31
3.7	Center pooling strategy	32
3.8	Coding dictionaries	32
3.9	Changes to protocol-defined analyses	32
4	STATISTICAL/ANALYTICAL ISSUES	32
4.1	Adjustments for covariates	32
4.2	Handling of dropouts or missing data	32
4.2.1	Primary efficacy variable	32
4.2.2	Secondary efficacy variables	33
4.2.3	Other efficacy variables	34
4.2.4	Safety variables	35
4.2.5	Other considerations	37
4.3	Interim analyses and data monitoring	37
4.4	Multicenter studies	38
4.5	Multiple comparisons/multiplicity	38
4.6	Use of an efficacy subset of subjects	38
4.7	Active-control studies intended to show equivalence	39
4.8	Examination of subgroups	39
5	STUDY POPULATION CHARACTERISTICS	40
5.1	Subject disposition	40
5.2	Protocol deviations	40
6	DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS	40
6.1	Demographics	41
6.2	Other Baseline characteristics	41
6.3	Medical history and concomitant diseases	43
6.4	Prior and concomitant medications	43

---

7	MEASUREMENTS OF TREATMENT COMPLIANCE.....	45
8	EFFICACY ANALYSES .....	46
8.1	Statistical analysis of the primary efficacy variable .....	47
8.1.1	Derivation of primary efficacy variable .....	48
8.1.2	Primary analysis of the primary efficacy variable .....	49
8.1.3	Supportive and sensitivity analyses of the primary efficacy variable .....	49
8.1.3.1	Including observed data at Week 52 .....	50
8.1.3.2	Multiple imputation .....	50
8.1.3.4	Observed case analysis .....	52
8.2	Statistical analysis of secondary efficacy variables .....	53
8.2.1	Derivation of secondary efficacy variables .....	53
8.2.1.1	Assessment in Axial SpondyloArthritis international Society response criteria (ASAS20/40, ASAS5/6, and ASAS partial remission).....	53
8.2.1.2	Bath Ankylosing Spondylitis Disease Index (BASDAI) .....	53
8.2.1.3	Bath Ankylosing Spondylitis Functional Index (BASFI) .....	54
8.2.1.4	Magnetic Resonance Imaging (MRI) assessments.....	54
8.2.1.5	Relevant changes to background medication .....	56
8.2.1.6	Ankylosing Spondylitis Quality of Life (ASQoL) .....	56
8.2.1.7	Nocturnal Spinal Pain.....	56
8.2.3.1	Multiple imputation of secondary continuous efficacy variables.....	59
8.3	Analysis of other efficacy variables during the double-blind 52-week treatment period	60
8.3.1	Derivation of other efficacy variables .....	62
8.3.1.1	Bath Ankylosing Spondylitis Metrology Index (BASMI) .....	62
8.3.1.2	Total spinal pain.....	63
8.3.1.3	SI joint radiographs .....	63
8.3.1.4	Tender and Swollen Joint Counts (44 joints evaluation).....	63
8.3.1.5	Patient's Global Assessment of Disease Activity (PGADA).....	64
8.3.1.6	Physician's Global Assessment of Disease Activity (PhGADA) .....	64
8.3.1.7	C-reactive protein (CRP).....	64
8.3.1.8	Spinal mobility .....	64
8.3.1.9	Enthesitis (MASES) .....	64
8.3.1.10	Short-Form 36-item Health Survey (SF-36) .....	65
8.3.1.11	ASAS-NSAID score.....	65
8.3.1.12	Work Productivity Survey (WPS).....	66
8.3.1.13	Medical Outcomes Study (MOS) sleep scale.....	68
8.3.1.14	Health Status (EQ-5D) .....	68
8.3.1.15	Resources Utilization .....	69
8.3.1.16	Extra-articular assessments .....	70

---

8.4	Analysis of other efficacy variables during SFE .....	70
9	PHARMACOKINETICS , PHARMACODYNAMICS, AND PHARMACOGENOMICS .....	70
9.1	Pharmacokinetics .....	70
9.2	Pharmacodynamics .....	71
9.3	Pharmacogenomics, biomarkers and cytokines .....	71
10	IMMUNOGENICITY .....	71
10.1	Available data .....	71
10.2	Subject classification .....	72
10.3	Analysis.....	73
11	SAFETY ANALYSES.....	74
11.1	Extent of exposure .....	74
11.2	Adverse events .....	75
11.3	Clinical laboratory evaluations .....	80
11.4	Vital signs, physical findings, and other observations related to safety .....	81
11.4.1	Vital signs .....	81
11.4.2	Electrocardiograms .....	82
11.4.3	Other safety variables .....	82
11.4.3.1	Weight .....	82
11.4.3.2	Pregnancy testing .....	82
11.4.3.3	Physical assessments.....	82
11.4.3.4	Tuberculosis assessments.....	82
12	REFERENCES .....	83
13	APPENDICES .....	84
13.1	Handling of questionnaire data .....	85
13.2	ASAS-NSAID equivalent score.....	86
14	AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN (SAP).....	87
	STATISTICAL ANALYSIS PLAN SIGNATURE PAGE.....	163

---

## LIST OF ABBREVIATIONS

ADAb	anti-drug antibody
ADaM	analysis data model
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AP	alkaline phosphatase
ASAS	Assessment of SpondyloArthritis international Society
ASAS20, 40	Assessment of SpondyloArthritis international Society 20%, 40% response criteria
ASAS5/6	Assessment of SpondyloArthritis international Society 20% improvement in 5 of 6 domains
AS	ankylosing spondylitis
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASDAS-CII	Ankylosing Spondylitis Disease Activity Score – Clinically Important Improvement
ASDAS-HD	Ankylosing Spondylitis Disease Activity Score – High Disease activity
ASDAS-ID	Ankylosing Spondylitis Disease Activity Score – Inactive Disease
ASDAS-LD	Ankylosing Spondylitis Disease Activity Score – Low Disease
ASDAS-MD	Ankylosing Spondylitis Disease Activity Score – Moderate Disease
ASDAS-MI	Ankylosing Spondylitis Disease Activity Score – Major Improvement
ASDAS-vHD	Ankylosing Spondylitis Disease Activity Score – very High Disease activity
ASQoL	Ankylosing Spondylitis Quality of Life

---

ASspiMRI-a	Ankylosing Spondylitis spine MRI acuity
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AU	anterior uveitis
AxSpA	axial spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASDAI50	Bath Ankylosing Spondylitis Disease Activity Index 50% response criteria
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BMI	body mass index
BMP	bone morphogenic protein
BP	bodily pain
CI	confidence interval
CK	creatinine kinase
COVID-19	Corona Virus Disease 2019
CR	compliance ratio
CRP	C-reactive protein
CSR	clinical study report
CV	coefficient of variation
CZP	certolizumab pegol
DEM	Data Evaluation Meeting
DKK1	dickkopf-related protein 1
DRL	Drug Reference List
DMARD	disease-modifying antirheumatic drug

---

EAIR	exposure-adjusted incidence rate
EAER	exposure-adjusted event rate
eCRF	electronic case report form
EQ-5D	EuroQoL Health Status Questionnaire (5 dimensions)
ER	emergency room
ES	Enrolled Set
FAS	Full Analysis Set
GH	general health
HCQ	hydroxychloroquine
HLA-B27	human leukocyte antigen B27
HLT	High Level Term
HRQoL	health-related quality of life
ICH	International Conference on Harmonisation
IXRS	interactive response system
IBD	inflammatory bowel disease
LOCF	last observation carried forward
LLN	lower limit of normal
LLOQ	lower limit of quantification
LLT	Low Level Term
MASES	Maastricht Ankylosing Spondylitis Enthesitis Score
MCS	Mental Component Summary
MedDRA®	Medical Dictionary of Regulatory Activities®
MH	mental health
MMP	matrix metalloproteinase

---

mNY	Modified New York (criteria)
MOS	Medical Outcomes Study
MRI	magnetic resonance imaging
MTX	methotrexate
nADA	neutralizing ADA <sub>b</sub>
nr-axSpA	nonradiographic axSpA
NRS	Numeric Rating Scale
NRI	non-responder imputation
NSAID	nonsteroidal anti-inflammatory drug
OC	observed case
OL	open-label
OL-CZP	open-label CZP
OL-OT	open-label other treatment
PBO	placebo
PCS	Physical Component Summary
PF	physical function
PGADA	Patient's Global Assessment of Disease Activity
PhGADA	Physician's Global Assessment of Disease Activity
PPS	Per Protocol Set
PT	Preferred Term
Q	question
Q1	1 <sup>st</sup> Quartile
Q3	3 <sup>rd</sup> Quartile
Q2W	every 2 weeks (every other week)

---

QoL	quality of life
RBC	red blood cell
RCTC	Rheumatology Common Toxicity Criteria
RE	role emotional
RP	role physical
RS	Randomized Set
SAARD	slow-acting anti-rheumatic drug
SAP	statistical analysis plan
sc	subcutaneous(ly)
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SF	social functioning
SFE	safety follow-up extension
SFE-WD	safety follow-up extension-early withdrawal visit
SFE-FAS	safety follow-up extension full analysis set
SF-36	Short-Form 36-Item Health Survey
SI	sacroiliac
SJC	swollen joint count
SM	spinal mobility
SMQ	Standardized MedDRA Query
SOC	System Organ Class
SOP	standard operating procedure
SpA	spondyloarthritis

---

SPARCC	Spondyloarthritis Research Consortium of Canada
sqrt	square root
SS	Safety Set
SSZ	sulfasalazine
STIR	short-tau-inversion recovery
TB	tuberculosis
TEAE	treatment-emergent adverse event
TJC	tender joint count
TNF	tumor necrosis factor
USA	United States of America
ULN	upper limit of normal
VAS	visual analog scale
VT	vitality
VU	vertebral units
WBC	white blood cell
WHO	World Health Organization
WPS	Work Productivity Survey

## 1 INTRODUCTION

This SAP describes the analysis of the double-blind 52-week treatment period and the 2 year open-label, Safety Follow-up Extension (SFE) of the C-axSpAnd (AS0006) study. It is designed to support a Clinical Study Report (CSR), is compliant with International Conference on Harmonisation (ICH) guidelines and is based on the Protocol Amendment 5 dated 17 December 2018.

## 2 PROTOCOL SUMMARY

The C-axSpAnd study (AS0006) is a randomized, double-blind, parallel-group, placebo-controlled study, consisting of 52 weeks of treatment with certolizumab pegol (CZP), followed by a follow-up period for 8 weeks after the Week 52/Withdrawal [WD] visit which is 10 weeks after the last administration of study medication, or at the completion of the Week 52 visit assessments, eligible subjects may receive open-label CZP treatment for an additional 2 years in an open-label Safety Follow Up Extension (SFE) Period.

The study population is subjects with active axial spondyloarthritis (axSpA) with sacroiliitis on magnetic resonance imaging (MRI) or C-reactive protein (CRP) levels indicative of inflammatory disease but without x-ray evidence of ankylosing spondylitis (AS) who have had an inadequate response to, have a contraindication to, or are intolerant to nonsteroidal anti-inflammatory drugs (NSAIDs).

Eligible subjects will be allocated to the following study treatments in a 1:1 ratio:

- CZP administered subcutaneously (sc) at the dose of CZP 400mg at Weeks 0, 2, and 4 followed by CZP 200mg every other week (Q2W) starting at Week 6
- Placebo

### 2.1 Study objectives

#### 2.1.1 Primary objective

The primary objective of the study is to demonstrate the efficacy of CZP 200mg Q2W on the signs and symptoms of subjects with active axSpA without x-ray evidence of ankylosing spondylitis.

#### 2.1.2 Secondary objectives

The secondary objectives of the study are to assess efficacy, safety, and tolerability and to demonstrate the effect of CZP on:

- Health outcomes
- Disease activity
- SI joint inflammation
- Changes to concomitant and background medications

#### 2.1.3 Other objectives

The other objectives of the study are to assess the effects of CZP on the following:

- Spinal mobility

- Total spinal pain numeric rating scale (NRS)
- Spinal inflammation
- SI joint structural changes
- Treatment response over time
- Additional signs and symptoms of the disease
  - Morning stiffness
  - Fatigue
  - Extra articular manifestations of axSpA
  - Sleep
  - Physical function
- Subject's health status
- Acute phase reactant (CRP)
- Health-related quality of life
- Work and household productivity
- Pharmacokinetics and immunogenicity
- Gene and protein expression and to explore the relationship between genomic biomarkers and disease biology, drug treatment and inflammatory and immune responses (for those subjects who consent to the genomics substudy)

## **2.2 Study variables**

### **2.2.1 Efficacy variables**

#### **2.2.1.1 Primary efficacy variable**

- ASDAS-MI response at Week 52

For Canada (and any other country where applicable or where requested by Regulatory Authorities), the primary efficacy variable is:

- Assessment of SpondyloArthritis international Society 40% response criteria (ASAS40) response at Week 12

#### **2.2.1.2 Secondary efficacy variables**

- ASAS40 response at Weeks 12 and 52
- Change from Baseline in BASFI at Weeks 12 and 52
- Change from Baseline in BASDAI at Weeks 12 and 52
- Change from Baseline in Sacroiliac (SI) joint SPARCC score at Week 12
- Number of subjects without relevant changes to background medication
- Change from Baseline in ASQoL at Week 52

- Change from Baseline in ASQoL at time points other than Week 52
- Change from Baseline in nocturnal spinal pain (NRS) at Week 52
- Number of subjects with Anterior Uveitis (AU) or new AU flares through Week 52

For Canada (and any other country where applicable or where requested by Regulatory Authorities), the secondary efficacy variables are:

- ASAS40 response at Week 52
- ASDAS-MI at Week 52
- Change from Baseline in BASFI at Weeks 12 and 52
- Change from Baseline in BASDAI at Weeks 12 and 52
- Change from Baseline in SI joint SPARCC score at Week 12
- Number of subjects without relevant changes to background medication
- Change from Baseline in ASQoL at Week 52
- Change from Baseline in ASQoL at time points other than Week 52
- Change from Baseline in nocturnal spinal pain (NRS) at Week 52
- Number of subjects with AU or new AU flares through Week 52

### 2.2.1.3 Other efficacy variables

The following variables will be analyzed at scheduled time points through Week 52 and at Week 156 for selected variables:

- ASAS 20% response criteria (ASAS20), ASAS40, ASAS 20% improvement in 5 of 6 domains (ASAS5/6), and ASAS partial remission response
- Change from Baseline in individual ASAS components:
  - Patient's Global Assessment of Disease Activity (PGADA)
  - Total spinal pain (NRS)
  - BASFI
  - Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
  - Change from Baseline in BASMI lateral spine flexion
  - CRP
- Change from Baseline in BASDAI and individual Questions 1, 2, 3 and 4
- Change from Baseline in ASDAS
- Change from Baseline in BASMI linear
- ASDAS disease activity (ASDAS-ID, ASDAS-LD\*, ASDAS-HD, ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI)
  - \* ASDAS-MD is equivalent to ASDAS-LD. Whereas protocol amendment 5 uses ASDAS-LD throughout the

SAP and Shells will continue to use ASDAS-MD in order to maintain consistency with the double-blind TFLs previously produced for this study.

- BASDAI50 response
- Change from Baseline in Fatigue (NRS) (from BASDAI)
- Change from Baseline in SI grading to Week 52 for structural damage
- Change from Baseline in SI joint SPARCC score at Week 52 and Week 156 and ASspiMRI-a in the Berlin modification at Week 12 and Week 52
- Proportion of subjects with SI joint SPARCC score <2 at Week 12, Week 52, and Week 156
- Change from Baseline in ASAS-NSAID score
- Number of AU flares
- Number of inflammatory bowel disease (IBD) exacerbations
- Number of psoriasis exacerbations
- Work Productivity Survey (WPS)
- Change from Baseline in the Sleep Problems Index II domains of the MOS Sleep scale
- Change from Baseline in enthesitis Maastricht Ankylosing Spondylitis Enthesitis Score (MASES)
- Change from Baseline in swollen and tender joint counts (44 joint count)
- Change from Baseline in Physician's Global Assessment of Disease Activity (PhGADA)
- Change from Baseline in the SF-36 PCS and MCS
- Change from Baseline in the SF-36 domains:
  - Role Physical
  - Bodily Pain
  - General Health
  - Vitality
  - Social Functioning
  - Role Emotional
  - Mental Health
- Health status as assessed by the EuroQoL Health Status Questionnaire (5 dimensions) (EQ-5D) domains, visual analog scale (VAS) actual score, and change from Baseline in VAS score
- Resources utilization: concomitant medical procedures, health care provider consultations not foreseen by the protocol, hospitalizations/emergency room visits

## **2.2.2 Pharmacokinetic, exploratory biomarker, and pharmacogenomics variables**

### **2.2.2.1 Primary Pharmacokinetic variables**

Certolizumab pegol plasma concentrations will be measured and summarized at Baseline and Weeks 1, 2, 4, 12, 24, 36, 52, and at the Follow-up visit (8 weeks after the Week 52/WD Visit).

These plasma samples may be used additionally for analyses of CZP and/or its constituent moieties, using alternative methods.

### **2.2.2.2 Other biomarker variables**

Selected samples collected for measurement of CZP plasma concentration may be used additionally for analyses of candidate biomarkers and cytokines, where appropriate. The biomarkers to be analyzed may include, but will not be limited to, the following:

Matrix metalloproteinase-3 (MMP-3), Bone morphogenic protein BMP-2, -4 and -7, wingless related mouse mammary tumor virus integration site protein (WNT1) - Inducible Signaling Pathway proteins (WISP), Gremlin, Dickkopf-related protein 1 (DKK1), Sclerostin, Hedgehog proteins (Shh, Ihh, Dhh), Collagen turnover/cleavage products, collagen type X, Vascular Endothelial Growth Factor (VEGF), citrullinated vimentin fragments, cytokines: IL13, IL17 A and F, IL23, IL34, Transforming Growth Factor (TGF)  $\beta$ , Macrophage colony-stimulating factors (M-CSF), Granulocyte macrophage colony-stimulating factor (GM-CS), colony-stimulating factor -1 (CSF-1), soluble CSF-1 Receptor (sCSF1r) levels.

### **2.2.2.3 Other pharmacogenomics variables**

For individuals consenting to the genomics substudy, blood samples will be drawn for possible genetic/epigenetic, genomic, proteomic, and metabolomics analysis at Baseline and Week 12. Additional samples will be collected for genomics, proteomics, and metabolomics analysis only, at Baseline, Weeks 4, 12, and 52/WD. Collection of the samples will enable exploratory evaluation of biomarkers relative to disease biology, drug treatment and inflammatory and immune response processes. The nature and format of these tentative analyses will be determined when the results of the main study are made available. The samples will be stored at -80°C at the central biorepository for up to 20 years.

### **2.2.3 Other immunological variables**

Anti-drug antibody (ADAb) (Anti-CZP antibody) concentrations will be assessed at Baseline and Weeks 1, 2, 4, 12, 24, 36, 52/WD, and the FU visit (8 weeks after the Week 52/WD visit). In addition, the ADAb will be assessed in subjects who withdraw from double-blind study drug and transition to open-label CZP prior to Week 52.

Determination of ADAb will be done using a validated screening, confirmation and titration ADAb bridging assay, with potential further characterization by a neutralizing antibody assay. The immunogenicity data will be processed according to dedicated Bioanalytical Analysis plans.

The following variables will be analyzed as described in Section 10:

- Anti-CZP antibodies at Baseline, Weeks 1, 2, 3, 12, 24, 36 and 52/WD
- Status of ADAb (including overall, baseline and treatment-emergent classification)

- ADA response characterized for their neutralizing potential

## **2.2.4 Safety variables**

### **2.2.4.1 Secondary safety variables**

#### **2.2.4.1.1 Adverse events**

Presence (Y/N) of

- Adverse Event (AE)
- Serious AE
- Non-serious AE
- Death
- AE leading to permanent withdrawal of study medication
- Severe AE
- Drug-related AE
- AE of special interest
  - Serious infections including opportunistic infections
  - Malignancies including lymphoma
  - Serious cardiovascular events
  - Congestive heart failure
  - Demyelinating-like disorders
  - Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia and leucopenia
  - Serious bleeding events
  - Lupus and lupus-like illness
  - Serious skin reactions (eg, Stevens Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme)

#### **2.2.4.2 Other safety variables**

##### **2.2.4.2.1 Laboratory parameters**

- Change from Baseline in hematology parameters at Weeks 2, 4, 8, 12, 24, 36, 52/WD and the Follow-up visit
  - Red blood cells
  - Hemoglobin
  - Hematocrit
  - Platelets
  - White blood cells

- 
- Neutrophils
  - Lymphocytes
  - Monocytes
  - Eosinophils
  - Basophils
  - Change from Baseline in hematology parameters to minimum post-Baseline value, maximum post-Baseline value, and last value
    - Red blood cells
    - Hemoglobin
    - Hematocrit
    - Platelets
    - White blood cells
    - Neutrophils
    - Lymphocytes
    - Monocytes
    - Eosinophils
    - Basophils
  - Change from Baseline in serum biochemistry parameters Weeks 2, 4, 8, 12, 24, 36, 52/WD, and the Follow-up visit
    - Sodium
    - Potassium
    - Chloride
    - Bicarbonate
    - Total calcium
    - Inorganic phosphorus
    - CRP (already included in efficacy section)
    - Creatinine kinase (CK)
    - Glucose
    - Creatinine
    - Uric acid
    - Urea
    - Total protein

- 
- Albumin
  - Alkaline phosphatase (AP)
  - Aspartate aminotransferase (AST)
  - Alanine aminotransferase (ALT)
  - Bilirubin
  - Total cholesterol
  - Change from Baseline in serum biochemistry parameters to minimum post-Baseline value, maximum post-Baseline value, and last value
    - Sodium
    - Potassium
    - Chloride
    - Bicarbonate
    - Total calcium
    - Inorganic phosphorus
    - CRP (already included in efficacy section)
    - Creatinine kinase (CK)
    - Glucose
    - Creatinine
    - Uric acid
    - Urea
    - Total protein
    - Albumin
    - Alkaline phosphatase (AP)
    - Aspartate aminotransferase (AST)
    - Alanine aminotransferase (ALT)
    - Bilirubin
    - Total cholesterol
  - Urinalysis status at Weeks 2, 4, 8, 12, 24, 36, 52/WD, and the Follow-up visit
    - PH
    - Protein
    - Glucose

- 
- Blood
  - Esterase
  - Hematology parameter normal range classification at Weeks 2, 4, 8, 12, 24, 36, 52/WD, and the Follow-up visit
    - Red blood cells
    - Hemoglobin
    - Hematocrit
    - Platelets
    - White blood cells
    - Neutrophils
    - Lymphocytes
    - Monocytes
    - Eosinophils
    - Basophils
  - Serum biochemistry parameter normal range classification at Weeks 2, 4, 8, 12, 24, 36, 52/WD, and the Follow-up visit
    - Sodium
    - Potassium
    - Chloride
    - Bicarbonate
    - Total calcium
    - Inorganic phosphorus
    - CRP (already included in efficacy section)
    - Creatinine kinase
    - Glucose
    - Creatinine
    - Uric acid
    - Urea
    - Total protein
    - Albumin
    - Alkaline phosphatase
    - Gamma glutamyl transferase

- 
- Aspartate aminotransferase
  - Alanine aminotransferase
  - Bilirubin
  - Total cholesterol
  - Hematology parameter marked abnormality classification (Rheumatology Common Toxicity Criteria [RCTC]) Weeks 2, 4, 8, 12, 24, 36, 52/WD, and the Follow-up visit
    - Hemoglobin
    - Platelets
    - White blood cells
    - Neutrophils
    - Lymphocytes
  - Serum biochemistry parameter marked abnormality classification (RCTC) Weeks 2, 4, 8, 12, 24, 36, 52/WD, and the Follow-up visit
    - Sodium
    - Potassium
    - Total calcium
    - Creatinine kinase
    - Glucose
    - Creatinine
    - Uric acid
    - Alkaline phosphatase
    - Aspartate aminotransferase
    - Alanine aminotransferase
    - Bilirubin
  - Presence (Y/N) of liver parameter elevations above upper limit of normal (ULN)
    - 3x ULN elevations of AST
    - 5x ULN elevations of AST
    - 10x ULN elevations of AST
    - 20x ULN elevations of AST
    - 3x ULN elevations of ALT
    - 5x ULN elevations of ALT
    - 10x ULN elevations of ALT

- 
- 20x ULN elevations of ALT
  - 3x ULN elevations of either AST or ALT
  - 5x ULN elevations of either AST or ALT
  - 10x ULN elevations of either AST or ALT
  - 20x ULN elevations of either AST or ALT
  - 1x ULN elevations of bilirubin
  - 1.5x ULN elevations of bilirubin
  - 1.5x ULN elevations of AP
  - 1x ULN elevation of bilirubin and 3x ULN elevation of either ALT or AST

#### **2.2.4.2.2 Vital signs**

- Change from Baseline in vital signs at Weeks 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52/WD, and the Follow-up visit
  - Pulse rate
  - Systolic blood pressure
  - Diastolic blood pressure
  - Temperature
  - Respiration rate
- Change from Baseline in vital signs to minimum post-Baseline value, maximum post-Baseline value, and last value
  - Pulse rate
  - Systolic blood pressure
  - Diastolic blood pressure
  - Temperature
  - Respiration rate

#### **2.2.4.2.3 Additional safety variables**

- Change from Baseline in weight at Weeks 12, 16, 24 and 52/WD
- Signs and symptom of latent or active TB completed at Baseline, Weeks 12, 24, 36 and 52/WD
- TB risk factors
- Occurrence of pregnancy through to Week 52/WD and Follow-up visit

## 2.3 Study design and conduct

The C-axSpAnd study (AS0006) is a 52-week multicenter, randomized, double-blind, parallel-group, placebo-controlled clinical study to evaluate the efficacy and safety of CZP in adult subjects with active axSpA with sacroiliitis on magnetic resonance imaging (MRI) and/or CRP levels indicative of inflammatory disease but without x-ray evidence of ankylosing spondylitis (AS) who have had an inadequate response to, have a contraindication to, or are intolerant to nonsteroidal anti-inflammatory drugs (NSAIDs). At the completion of the Week 52 visit assessments, eligible subjects may receive open-label CZP treatment for an additional 2 years in an open-label, Safety Follow-up Extension (SFE) period.

### 2.3.1 Study periods

Period 1 (Screening period) – Screening period of 1 day to 6 weeks before Baseline in order to obtain laboratory data, to verify that the doses of methotrexate (MTX), sulfasalazine (SSZ), hydroxychloroquine (HCQ), NSAIDs, and corticosteroids, if used, are stable, and to enable washout of any medications not permitted for use during the study, and initiation of latent TB treatment where necessary. Concomitant medications permitted during the study should remain stable for at least 4 weeks prior to Baseline at Week 0.

SI-joint x-rays (read centrally) will allow discrimination of subjects with (AS) and without definitive evidence for sacroiliitis on x-ray (modified New York [mNY]-negative-axSpA).

Enrolled subjects must undergo an MRI later during the Screening period for central reading with results from the central reading available by no later than at the Baseline visit.

Period 2 (Double-blind period) – Week 0 to Week 52, placebo-controlled.

Eligible subjects will be allocated to the following double-blind study treatments in a 1:1 ratio:

- CZP administered sc at the loading dose of CZP 400mg Q2W at Weeks 0, 2 and 4 followed by CZP 200mg Q2W (starting at Week 6)
- Placebo

At on-site visit days (at Weeks 0, 2, 4, 6, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 50), the double-blind study treatments (including placebo) will be administered sc at the study site. At the Week 0, 2, 4, 6, and 8 visits, the administration must be done by dedicated, unblinded, and adequately trained site personnel. At the Week 10 and 12 visits, the subject will receive training on self-injection and the study treatment will be self-administered under the supervision of the unblinded study personnel. At Weeks 16, 20, 24, 28, 32, 36, 40, 44, 48, and 50, the study treatment administration should preferably be done by the unblinded study personnel, or alternatively by self-injection of the subject under the supervision of the unblinded study personnel.

Subjects who discontinue the double-blind study treatment and enter open-label treatment with CZP will be assessed 2 and 4 weeks after open-label CZP treatment has been initiated and every 12 weeks thereafter. Alternatively, subjects who discontinue the double-blind study treatment and receive other treatment (including biologics) will be assessed every 12 weeks after initiation of the open-label other treatment (OL-OT). In either alternative schedule, assessments should be conducted until as close as possible to Week 52 (within  $\pm 4$  weeks), where Week 52 is relative to

the original randomization at Week 0. Subjects will then be invited to the Week 52 assessment visit.

**Period 3** (Follow-up period) - All subjects not participating in the SFE period after Week 52, including those withdrawn from the study prematurely, will have a single Follow-up visit, 8 weeks after Week 52/WD visit (10 weeks after their last administration of study medication).

**SFE Period** - Week 52 to Week 156, open-label:

At the completion of the Week 52 visit assessment, subjects on double-blind study treatment may receive open-label CZP treatment for an additional 2 years. Subjects who withdrew from double-blind study treatment and transitioned to open-label CZP treatment are also eligible to participate in the SFE period after completing the Week 52 visit assessment. Subjects on OL-OT are not eligible to participate in the SFE period.

Eligible subjects are allowed to roll-over to the SFE-period up to 3 months after completion of the Week 52 assessments.

Starting at Week 64, all subjects are to visit the site approximately every 12 weeks ( $\pm 2$  weeks) for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156/Safety Follow-up Extension-Early Withdrawal Visit (SFE-WD) ( $\pm 1$  week). However a  $\pm 4$  week window will be used in the analysis for the Week 156 timepoint. The expected day for Week 156 will be calculated assuming that the SFE period start was on Week 52, i.e. Week 156 will be expected to be  $(156-52) \times 7$  days after Week 52. For patients where the first dose after the Week 50 visit is before the SFE informed consent the width of the lower boundary of the Week 156 window will be increased further by the absolute number of days between the first dose after the Week 50 visit and the SFE informed consent date.

The schedule of study assessments is provided in [Section 5.2](#) of the protocol.

A schematic diagram of the study design is provided in [Section 5.3](#) of the protocol.

### **2.3.2 Study duration per subject**

For each subject, the first 3 periods of the study will last a maximum of 66 weeks, as follows:

- Up to 6 weeks in the Screening period
- 52 weeks in the double-blind period
- A Follow-up visit, 8 weeks after Week 52/WD visit for subjects not participating in the SFE Period.

For subjects participating in the SFE period, the study will extend up to a maximum of 104 additional weeks.

The end of the study will be defined as the date of the last subject's last visit, defined as the Follow-up Visit 8 weeks after Week 52/WD visit for subjects not participating in the SFE Period, or the last visit of the SFE Period (Week 156/SFE-WD).

### **2.3.3 End of periods and study**

The end of the study is defined as the date of the last visit (including Follow-up visit and SFE period) of the last subject in the study. Period start and ends are defined in [Section 3.2.1](#).

### **2.3.4 Planned number of subjects and sites**

Approximately 1200 subjects are expected to enter the Screening period in order to have 300 subjects randomized into the double-blind period. It is planned to enroll the subjects at approximately 120 sites.

### **2.3.5 Anticipated regions and countries**

The study will be conducted in North America, Australia, Europe, Asia, and other regions as appropriate.

### **2.3.6 Stratification**

Randomization will be stratified on:

- Region
  - North America: Canada and USA
  - Europe: Bulgaria, Czech Republic, Hungary, Poland and Russia
  - Asia: Australia, and Taiwan
- MRI/CRP classification

Subjects will be classified as MRI+/- depending on whether or not they have evidence of sacroiliitis on MRI at Screening based on the ASAS/OMERACT definition. Subjects will be classified as CRP+/- based on the CRP value obtained at the second Screening visit scheduled to occur 3 to 5 days prior to Baseline. Subjects will be categorized as CRP+ if their CRP value is above the level indicative of inflammatory disease at this visit. Otherwise, they will be considered CRP-. Based on these definitions, the stratification for MRI/CRP classification will have the following 3 levels:

- MRI+/CRP+
- MRI+/CRP-
- MRI-/CRP+

Subjects who are MRI-/CRP- are not eligible for randomization. The interactive response system (IXRS) will be designed to ensure that at least 20% of the randomized subjects belong to one of the three clinical subgroups above.

## **2.4 Determination of sample size**

Subjects will be randomized in a 1:1 ratio to the CZP 200mg Q2W and placebo treatment groups. The expected response rates for ASDAS-MI at Week 52 are 40% and 20% for CZP and placebo, respectively. A total sample size of 300 (150 subjects per treatment group) provides 95% power to detect a statistically significant difference in the ASDAS-MI response rate at Week 52 between CZP and placebo based on a 2-sided significance level of 0.05.

## **3 DATA ANALYSIS CONSIDERATIONS**

### **3.1 General presentation of summaries and analyses**

All statistical analyses will be performed using SAS® (STATISTICAL ANALYSIS SYSTEM, SAS-Institute, Cary, NC, USA) according to UCB SOPs.

For continuous data in general, summary statistics (n [number of available measurements], arithmetic mean, SD, median, minimum, and maximum) will be presented by treatment group. For selected variables, the first (25<sup>th</sup> percent) quartile (Q1) and the third (75<sup>th</sup> 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 electronic case report form (eCRF) 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. P-values will be presented to 3 decimal places. Relevant SAS output will be included in the 'Documentation of Statistical Methods' section of the CSR. All data in the database (SDTM and ADaM) will be presented in by-subject data listings, and sorted by treatment group, site, subject number, and visit (where applicable).

A fixed-sequence testing procedure will be used to control the overall Type 1 error for comparison of multiple efficacy endpoints. This procedure is described in [Section 8](#).

## **3.2 General study level definitions**

### **3.2.1 Analysis time points**

#### **3.2.1.1 Pre-treatment period**

The pre-treatment period (Screening period) of the study is the period prior to a subject's first dose of study medication administration. This period starts at the Screening visit (week -6 to -1 day) 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, all assessments will be attributed to the Screening period.

### 3.2.1.2 Double-blind treatment period

The double-blind treatment period begins at the Baseline visit (Week 0) at the time of first study medication administration (inclusive) and ends when:

- the subject takes their first dose of open-label CZP (OL-CZP) or open-label other treatment (OL-OT) for those who continue into the SFE or who switch to OL-CZP or OL-OT
- the later of Week 52 visit and last administration of double-blind treatment, for subjects who complete the 52 week main phase of the study and do not continue into the SFE and who do not switch to OL-CZP or OL-OT
- the subject discontinues the study on double-blind treatment prior to the week 52 visit (WD visit) and does not fall into the above categories.

For subjects who prematurely withdraw from the study prior to Week 52 (for example, those who cease to have double-blind and open-label visit data collected), their visit assessments at withdrawal will be assigned to the next scheduled visit following the last visit where assessments are available according to each protocol activity. This could be the next double-blind or open-label visit. 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 main phase of the study if they complete the Week 52 visit without early withdrawal during the main phase of the study. This is regardless of whether they attend the Follow-up visit or enter SFE period and regardless of whether they are on double-blind treatment, OL-CZP, or OL-OT.

### 3.2.1.3 Open-label period

The open-label period starts when subjects have discontinued the double-blind treatment period prior to start of the SFE period and take their first dose of open-label CZP or other treatment. The open-label period ends when:

- the subject takes their first dose of OL-CZP in the SFE period, for subjects continuing into the SFE
- the later of Week 52 visit and last administration of OL-CZP or other treatment, for subjects who complete the 52 week main phase of the study and do not continue into the SFE
- the subject discontinues the study prior to the week 52 visit (WD visit) and does not fall into the above categories.

Subjects switching from double-blind treatment to OL-CZP treatment and then to OL-OT will have an OL-CZP and an OL-OT period.

Subjects continuing on the study with other treatment visits, but without taking other medication will be allocated to the OL-OT period although will not be considered as starting other treatment in that period. A medical review of all cases will be performed prior to the database lock to identify subjects with valid administration of other treatment. Start and end dates of identified other treatments will be taken from the concomitant medication page.

The use of open-label period data for efficacy is described in detail in the relevant efficacy sections in [Section 8](#). If applicable, open-label period data for efficacy will be mapped to the corresponding time-point post baseline for analysis. E.g. if a subject performs an open-label

Week 12 visit at 26 weeks post-baseline, corresponding efficacy data will be considered “week 26” assessments.

Open-label visit-based safety data (such as laboratory data and vital signs data) will in general be summarized and listed as separate visits from the double-blind period visits. However, open-label visits will still be presented by the double-blind treatment groups of CZP 200mg Q2W and placebo. In general, OL-OT safety data will not be presented in tables, but only listed.

For subjects who receive OL-CZP, data will be presented by the OL-CZP visits. Safety summaries which are summarized ‘At any visit’, ‘post-Baseline’ (including minimum or maximum post-Baseline) or ‘On treatment’ will be presented separately for each period (double-blind and OL-CZP). Visit based efficacy or safety data that is recorded on the day of first administration of OL-CZP or OL-OT will be assigned to the previous period unless available time information clearly indicates that it was recorded after the first administration of OL-CZP or OL-OT.

Concomitant medications will be assigned to all periods (double-blind, OL-CZP, OL-OT or SFE) according to the rules given in [Section 6.4](#).

AEs will be assigned to a period based on whether the AE started during the double-blind, OL-CZP, OL-OT period or SFE period. See [Section 11.2](#) for more detail.

All data reported will be listed.

#### **3.2.1.4 Follow-up period**

For subjects not entering the SFE period (no SFE informed consent), the Follow-up period will start after the later of Week 52 visit or last study medication administration (for subjects having Week 52 visit on double-blind or open-label therapy) or after the WD visit. Visit based efficacy or safety data that is recorded at the Week 52 or WD visit will be assigned to the previous period.

The Follow-up period will end on the Follow-up visit date. The Follow-up visit will take place 10 weeks after the last dose of study medication, which should be 8 weeks after the Week 52 visit (if the Week 52 is the last non-Follow-up visit) or less than 10 weeks after the WD visit (if the WD visit is the last non-Follow-up visit).

#### **3.2.1.5 Pre-SFE period**

The pre-SFE period starts after last dose of the double-blind period (i.e. also accounts for the last OL dose if the subject escaped) + 70 days (note that the last double-blind dose is at Week 50, thus each AE which starts up to 70 days after the last double-blind dose is still assigned to the double-blind phase) and ends before SFE period starts. The first SFE dose is used to determine if AEs end before the SFE period starts since this is the time when the AEs become treatment-emergent.

#### **3.2.1.6 SFE period**

The SFE period starts when the subject takes their first dose of OL-CZP that is after week 50 and after date of informed consent for the SFE. Visit based efficacy or safety data that is recorded on the day of the first SFE period dose will be assigned to the previous period unless available time information clearly indicates that it was recorded after the first administration of SFE period treatment. There are some subjects who took their first dose of OL-CZP that is after week 50

before the informed consent for the SFE. In such cases the SFE period will be programmed to start at the informed consent date.

The SFE period ends when the subject discontinues from or completes the study following SFE enrollment (Week 156/SFE-WD).

### 3.2.2 Relative day

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

- If the start (stop) date occurred on or after the first dose, and on or prior to the double-blind drug stop date, relative day is calculated as start (stop) date minus first double-blind dose date + 1. Relative day will be prefixed by a 'd' in the listings to show it's relative to double-blind treatment.
- For subjects not entering the OL-CZP period (including SFE period), if the start (stop) date occurred after the last dose of double-blind drug, the relative day to the most recent double-blind dose is calculated as start (stop) date minus most recent double-blind dose date. The relative day in this situation should be preceded by a 'd+'.
- For subjects entering the OL-CZP period (including SFE period), if the start (stop) date occurred on or after the first OL-CZP start date, and on or prior to the OL-CZP stop date (including SFE period), relative day is calculated as start (stop) date minus first open-label dose date + 1. The relative day in this situation should be preceded by a 'o'.
- For subjects entering the OL-CZP period (including SFE period), if the start (stop) date occurred after the last dose of OL-CZP drug, the relative day to the most recent OL-CZP dose is calculated as start (stop) date minus most recent OL-CZP dose date. The relative day in this situation should be preceded by a 'o+'.
- 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 'd-'.

For AEs, relative days for start and stop dates will be calculated as the number of days since the first injection of the relevant double-blind or OL-CZP medication. For non-treatment emergent AEs, relative day of onset will be negative if the event started and stopped before the first dose.

### 3.3 Definition of Baseline values

Unless otherwise specified, the last valid measurement before study medication administration in the double-blind period will be used as the Baseline value. The same Baseline definition will be used for the open-label period and SFE period. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables (eg, demography) assessments are scheduled for the Screening visit only and not for the Baseline visit. In this case, the Screening value will be utilized as Baseline value. 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.

Baseline images, either spinal x-ray or MRI assessments, are required to evaluate study eligibility. Therefore, all Baseline images will be available prior to randomization into the study. SI joint x-rays should not be older than 12 months prior to Baseline and should be verified by central reading during the Screening period.

There will be an additional SFE Baseline presented in some tables. It is defined as the last value before the patient received SFE treatment (the first dose on or after SFE informed consent, for most patients this will be the Week 52 value).

### **3.4 Protocol deviations**

Important protocol deviations are deviations from the protocol which could potentially have a meaningful impact on study conduct or on the primary efficacy outcome for an individual subject. The criteria for identifying important protocol deviations will be defined within the specification of Protocol Deviations 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 unblinding to confirm exclusion from analysis sets.

### **3.5 Analysis sets**

#### **3.5.1 Enrolled Set**

The Enrolled Set (ES) will consist of all subjects who have given informed consent.

#### **3.5.2 Randomized Set**

The Randomized Set (RS) will consist of all subjects randomized into the study.

#### **3.5.3 Safety Set**

The Safety Set (SS) will consist of all subjects in the RS who have received at least one dose of study medication.

#### **3.5.4 Full Analysis Set**

The Full Analysis Set (FAS) will consist of all subjects in the RS who have received at least 1 dose of study medication.

#### **3.5.5 Per Protocol Set**

The Per Protocol Set (PPS) will consist of subjects in the FAS without any important protocol deviations that may influence the validity of the primary efficacy data. Treatment compliance as defined in Section 7 may also be utilized. Important protocol deviations will be predefined and evaluated at the DEM prior to study unblinding at the Week 52 interim database lock. Protocol deviations occurring after Week 52 (for example, for subjects continuing in the SFE period) will not be considered for PPS impact as they occurred after assessment of the primary variable was performed.

#### **3.5.6 SFE Safety Set**

The SFE Safety Set (SFE-SS) will be defined as all subjects who continued into the SFE Period and who received at least 1 dose of CZP in the SFE Period. Safety data recorded during the SFE will be presented alongside the Week 52 analyses and hence it is not expected that this analysis set will be required.

### 3.5.7 SFE Full Analysis Set

The SFE Full Analysis Set (SFE-FAS) will be defined as all subjects in the FAS, who continued into the SFE Period and who received at least 1 dose of CZP in the SFE Period. This analysis set will be used to analyze efficacy data at Week 156/SFE-WD.

### 3.6 Treatment assignment and treatment groups

At Baseline, eligible subjects will be randomly assigned in a 1:1 ratio to the following double-blind study treatments:

- Placebo (PBO)
- CZP administered sc at the loading dose of CZP 400mg Q2W at Weeks 0, 2 and 4 followed by CZP 200mg Q2W (starting at Week 6)

For presentation purposes, the terms CZP 200mg Q2W and PBO will be used.

For disposition, demography, Baseline characteristics and protocol deviations, in addition to the CZP 200mg Q2W and PBO groups, an all subjects group (CZP 200mg and PBO combined) will be displayed.

In general, for by visit data collected during the open-label period, subjects will be presented according to the treatments received in the double-blind period (CZP 200mg Q2W and PBO). For the by visit data collected during the SFE period, subjects will be presented in a single CZP 200mg Q2W treatment group.

Concomitant medications will be summarized according to whether they were taken: during the double-blind period by CZP 200mg Q2W or PBO, or during the open-label period by PBO->OL CZP (subjects who were randomized to PBO and then received open-label CZP), CZP->OL CZP (subjects who were randomized to CZP and then received open-label CZP)- or by the OL-OT group. In addition, concomitant medications taken during any CZP medication period of the main study (Total CZP incl DB and OL) and all concomitant medications (All subjects, including OL-OT but excluding SFE) will also be summarized. See [Section 6.4](#) for more detail.

AEs will be assigned to periods based on whether they started during the double-blind, OL-CZP, OL-OT, or SFE periods. Data corresponding to the OL-OT group will not be summarized, but only presented in listings. AEs will be summarized as follows:

- Double-blind period by CZP 200mg Q2W or PBO,
- Open-label CZP period by PBO->OL CZP or CZP->OL CZP (as above),
- SFE period by OL CZP
- Total CZP incl DB and OL (includes randomized CZP 200mg Q2W and AEs starting in the open-label CZP period).
- Total CZP incl DB, OL, and SFE OL (includes randomized CZP 200mg Q2W, AEs starting in the open-label CZP period and AEs starting in the SFE period).

See [Section 11.2](#) for more detail.

If it is determined after unblinding that the subject received a treatment other than the one to which they were randomized, safety tables will still be based on the randomized treatment for the SS. However, additional safety tables based on the actual treatment received may also be prepared whereby all data taken on or after the 1<sup>st</sup> administration of CZP is summarized according to the CZP treatment group no matter which treatment they were randomized to. The efficacy analyses will strictly follow the intention to treat principle, and no correction for receiving incorrect treatment will be performed.

### **3.7 Center pooling strategy**

Due to the anticipated low number of subjects per center, region is used as a randomization stratification factor. The regions utilized in the IXRS system (as described in [Section 2.3.6](#)) will be used to pool centers for statistical analysis. Verification that sufficient subjects within each region exist for meaningful statistical analysis will be performed prior to unblinding.

### **3.8 Coding dictionaries**

Medical history and AEs will be coded using the latest version available for use at the time of study lock of Medical Dictionary for Regulatory Activities (MedDRA). Medications will be coded using the World Health Organization Drug Reference List (WHO-DRL).

### **3.9 Changes to protocol-defined analyses**

The protocol stated that ADAb concentrations above 2.4 unit/ml would be defined as ADAb positive. However, during the trial it was noted that the background sample noise could only be determined during the sample analyses and hence the cut point cannot be pre-specified. Therefore the immunogenicity sections of the SAP were updated.

The protocol defines a plus/minus one week window for the Week 156 time point. However at least a plus/minus 4 week window will be used in the analysis due to the Corona Virus Disease 2019 (COVID-19) pandemic, with an even more generous lower boundary to the window for patients who had their first dose post-Week 50 before the SFE informed consent date.

## **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**

#### **4.2.1 Primary efficacy variable**

For analysis purposes, the primary efficacy variable is defined as a composite endpoint in which it is necessary to achieve ASDAS-MI at Week 52 and to remain on randomized double-blind study treatment through Week 52 in order to be considered a responder ([Section 8.1](#)). Subjects with missing ASDAS-MI at Week 52 are considered non-responders and therefore this composite endpoint definition does not allow for a missing response status and no formal method for handling missing data is needed in the primary efficacy analysis. However, in order to assess the impact of various missing data assumptions on the analysis, additional sensitivity analyses of the primary efficacy variable will be performed as follows:

- Including observed data at Week 52: The observed data at Week 52 will be used and analyzed according to randomized treatment, regardless of whether or not the subject is still on his/her randomized double-blind study treatment at that time. Therefore, this is no longer a composite endpoint and subjects will be a responder based solely on if they achieve ASDAS-MI response. Despite efforts to continue to collect data on all subjects (even if they discontinue the double-blind study treatment), there may still be subjects with missing data at Week 52. Such subjects will be considered as non-responders. The same logistic regression model specified for the primary analysis will be used.
- Multiple imputation: The MCMC method will be used to impute intermittent missing data. The resulting multiply imputed data sets will be monotone missing and will be imputed using monotone regression (assuming a MAR pattern). Note that the multiple imputation procedure will be done on the continuous ASDAS variable, which will be dichotomized for the logistic regression analysis. ASDAS data collected after the discontinuation of double-blind study treatment will be treated as missing in the original data set for this sensitivity analysis and hence will not be used for analysis or imputing values.
- Tipping point analysis: In this analysis, various assumptions will be made about average outcomes among the subsets of subjects who prematurely discontinued double-blind study treatment. Various “delta adjustments” will be made to the assumed responses among missing data in each treatment arm with varying degrees of plausibility (O’Kelly, 2014). The goal of the tipping point analysis is to identify assumptions about the missing data under which the conclusions change, i.e., under which there is no longer evidence of a treatment effect. Then, the plausibility of such assumptions would be discussed. Further details of this procedure are described in [Section 8.1.3.3](#).
- Observed case analysis: This analysis will only include the observed data for subjects still on the original double-blind study treatment. Data collected after the discontinuation of double-blind study treatment and all other missing data will be excluded from the analysis. The same logistic regression model specified for the primary efficacy analysis will be performed.

For Canada (and any other country where applicable or where requested by Regulatory Authorities), the primary efficacy variable is ASAS40 at Week 12 instead of ASDAS-MI at Week 52. For this reason, the above analyses will be repeated for ASAS40 at Week 12 with the composite endpoint analysis being the primary analyses.

#### 4.2.2 Secondary efficacy variables

ASAS40 at Week 12, although a secondary efficacy variable for the USA, is described within [Section 4.2.1](#) due to its status as a primary efficacy variable in other parts of the world.

The primary analysis method for the binary secondary efficacy variable(s), ASAS40 at Week 52, will be the same as that described above for the primary efficacy variable, ASDAS-MI. That is, a subject will be considered a responder only if the double-blind treatment course is completed and if ASAS40 is achieved. The sensitivity analyses of the secondary binary efficacy variable(s) (ASAS40 at Week 52 as described in [Section 8.2](#)) will be as follows:

- Including observed data at Week 52 (or Week 12 as applicable): As described for the primary efficacy variable.

- Multiple imputation: As described for the primary efficacy variable. Note that as ASAS40 is a composite of 4 different variables, the multiple imputation procedure will be performed on each of these components, and the ASAS40 response will be derived based on the multiply imputed datasets.
- Reference-based multiple imputation: This procedure will impute missing data as well as data collected at time points after the discontinuation of the double-blind study treatment for both the CZP and placebo groups. This procedure will use an imputation model developed based on data from the placebo group (Mallinckrodt, 2013). The estimand being evaluated in this case is the difference in the outcome of improvement in all randomized subjects at the planned endpoint, attributable to the initially randomized medication, which has been referred to as “estimand 6” (Mallinckrodt, 2012).
- Observed case analysis: As described for the primary efficacy variable.

For the continuous secondary efficacy variables (BASFI and BASDAI at Week 52 and 12, and SI joint SPARCC score at Week 12), the primary analysis method will be based on a referenced based multiple imputation procedure (details described in [Section 8.2](#)). The following sensitivity analyses will also be performed

- Including observed data at Week 52 (or Week 12 as applicable): As described for the primary efficacy variable
- Multiple imputation: As described for the primary efficacy variable.
- Observed case analysis: As described for the primary efficacy variable
- Last Observation Carried Forward (LOCF): If there is missing data at the double-blind time point of interest, the last available post-Baseline measurement will be carried forward to that time point (including any previous scheduled or unscheduled visits in double-blind or open-label periods). Using the date that the observation was recorded, data will be carried forward to all missing double-blind visits up until the time point of interest even if the subject has withdrawn from the study or entered the open-label period. For example, a subject withdrawing from double-blind at Visit 14 /Week 24 and entering OL-CZP will proceed to have a open-label visits at weeks 0, 2 and 4 post double-blind withdrawal followed by assessments every 12 weeks until 52 weeks post randomization. Using the date of the open-label visits, the number of weeks post randomization will be calculated to allow data to be carried forward to all subsequent missing double-blind visits. This will not apply for the analysis of SI joint SPARCC score at Week 12 since there are no measurements collected between Baseline and Week 12.

There will be no imputation or other method for handling missing data for the secondary efficacy variable, number of subjects without relevant changes to background medication. Further details on the derivation and analysis of this variable are described in [Section 8.2](#).

#### **4.2.3 Other efficacy variables**

For the analysis of the double-blind 52-week treatment period, other binary efficacy variables will be treated in the same way as the primary analysis of the primary efficacy variable (composite endpoint). That is, a subject will be considered a responder only if the response is achieved and if the subject is still on their randomized double-blind study treatment at the time

when the variable is evaluated. Analyses of other continuous efficacy variables will be based on LOCF. A sensitivity analyses will be performed on all other efficacy variables using observed case analysis as described for the primary endpoint in [Section 4.2.1](#).

For analysis of the Week 156/SFE-WD visit data, binary efficacy variables will be analyzed using the non-responder imputation (NRI) approach; that is, where the denominator is all subjects in the SFE-FAS for the given treatment group. A subject will be considered a responder only if the response criterion of the given variable is achieved. All other subjects will be counted as not responding e.g. subjects who have not achieved the given outcome and all subjects with missing data are set to non-response.

An observed case analysis of binary efficacy variables at Week 156/SFE-WD will be performed. The Observed Case analysis at Week 156, will consider a subject a responder only if the response is achieved. Non-response at Week 156 visit will be counted as non-responders. Early withdrawals or subjects with missing data will be considered missing in this analysis. The sensitivity analysis of the Observed Case analysis at Week 156/SFE-WD will consider a subject as a responder if response is achieved at Week 156/SFE-WD visit such that both subjects with a Week 156 visit (subjects who complete) and subjects with early withdrawal visits and efficacy results meeting the criteria of response, will be counted as a responder. Subjects with missing data will be considered missing in this analysis.

For continuous efficacy variables, the observed data at Week 156/SFE-WD visit will be analyzed. This will summarize efficacy data for completers and subjects with early withdrawal visits during SFE together. In addition, the efficacy data collected at Week 156 for completers only will be summarized.

#### 4.2.4 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. If time of study treatment and time of event or concomitant medication is available and non partial then it will also be used however if missing or partial, only the dates will be compared. 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.

- Missing start day, but month and year present:

If the start of double-blind 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 administration of study medication during the double-blind period. Otherwise if informed consent was signed 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 informed consent.

Otherwise if the following conditions are met:

- the year of the occurrence of the AE/concomitant medication is after the year of the end of the double-blind period or

- the year of the occurrence of the AE/concomitant medication is the same as the year of the end of the double-blind period and the month of the occurrence of the AE/concomitant medication is the after the month of the end of the double-blind period

then the following imputations will be done:

If the year of the occurrence of the AE/concomitant medication is the same as the start of the open-label period and the month of the AE/concomitant medication is the same as the month of the start of the open-label period then the start day of the event/concomitant medication will be assigned to the first day of the open-label period.

Otherwise if the year of the occurrence of the AE/concomitant medication is the same as the start of the SFE period and the month of the AE/concomitant medication is the same as the month of the start of the SFE period then the start day of the event/concomitant medication will be assigned to the first day of the SFE period.

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 double-blind 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 administration of double-blind study medication. Otherwise if informed consent was signed in the same year as the occurrence of the AE/concomitant medication, the start day and month will be assigned to the day and month of informed consent.

Otherwise if the year of the occurrence of the AE/concomitant medication is after the year of the end of the double-blind period and then the following imputations will be done:

If the year of the occurrence of the AE/concomitant medication is the same as the start of the open-label period then the start day and month of the event/concomitant medication will be assigned to the first day of the open-label period.

Otherwise if the year of the occurrence of the AE/concomitant medication is the same as the start of the SFE period then the start day of the event/concomitant medication will be assigned to the first day of the SFE period.

Otherwise the start day and month will be set to January 1st.

- Missing start day, month, and year:

The start date will be replaced by the date of first study drug administration of double-blind study medication or by the first day of January of the year of the end date, if the first administration date of double-blind study medication 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 administration 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 administration of study medication. Otherwise the end day will be set to December 31st of the given year.

- If the stop date is completely unknown, the stop date will not be imputed.

#### **4.2.5 Other considerations**

Descriptive summaries for all efficacy variables collected during the double-blind 52-week treatment period based on observed case data (including only subjects still on their randomized double-blind study treatment) will also be prepared.

### **4.3 Interim analyses and data monitoring**

An interim analysis is planned after the completion of the double-blind period of the last subject at Week 52. The timing of the analysis will be when the last subject completes Week 52 visit in accordance with the double-blind, OL-CZP or OL-OT schedule. At this time, all visit based data from the double-blind period up to and including Week 52 visit, all open-label visit based data up to Week 52 visit, all available Follow-up data for subjects already completing the study, all hospitalization/emergency room visit and health care provider consultation data, all study medication discontinuation and all prior and concomitant medication data including concomitant medical procedures, will be locked. Subjects still on the study awaiting Follow-up visit or in SFE, will continue to have AEs, laboratory, vital sign, PK and physical examination data collected per schedule until completion of the Follow-up or SFE period, at which point they will complete the study termination eCRF page. At the time of the interim analysis Week 52 DB lock, adverse event data entered to date will be cleaned but not locked. This will allow ongoing adverse event entry during the Follow-up and SFE periods. The treatment codes will be made available to the study reporting team and an interim study report will be written. All tables, listings and figures described in this SAP for analysis of the double-blind 52-week treatment period will be produced at the time of the Week 52 DB lock. The Investigators and subjects will remain blind to the assigned CZP dose regimen of the double-blind period.

After the last subject has completed the SFE period, the database will be fully locked, and a final study report will be written. For subjects in the SFE period, from the Week 52 visit onward, subjects will be treated with OL-CZP until Week 156/SFE-WD visit. During this time, only adverse events, concomitant medication related to adverse events, and study drug dispensation information will be collected. Efficacy assessments and study termination data will be collected at Week 156/SFE-WD as described in Table 5-6 of the study protocol. Therefore, at the final DB lock, the following tables and listings will be produced.

- Table 1.1.2: Disposition of Subjects Screened
- Table 1.2.1: Disposition and Discontinuation Reasons
- Table 1.2.2: Number of Subjects Completing each Visit
- Table 1.3: Disposition of Analysis Sets
- Table 1.4.1: Important Protocol Deviations During SFE
- Efficacy tables for analysis of efficacy data collected at Week 156/SFE-WD visit per Table 5-6 of the study protocol
- Table 7.1: Extent of Exposure

- Adverse Event Tables
- Laboratory Tables in order to include any data collected at the Follow-up visit post Week 52 interim database lock for subjects not entering SFE
- Listing 1.2: Subject Disposition
- Listing 1.3: Study Termination
- Listing 1.4: Visit Dates
- Listing 3.1: Important Protocol Deviations
- Listing 4.1: Subject Analysis Sets
- Efficacy listings for all efficacy data collected at Week 156/SFE-WD visit per Table 5-6 of the study protocol
- Adverse Event listings
- Laboratory listings in order to include any data collected at the Follow-up visit post Week 52 interim database lock for subjects not entering SFE.

Regular monitoring of safety data collected during the study 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.

#### **4.4 Multicenter studies**

Due to the low number of subjects expected to be recruited per site, results will not be tabulated by individual sites. However important protocol deviations will be tabulated by region and the number of subjects randomized per region will also be presented. Listings will show the region and site for each subject. In addition, as subject randomization is stratified by region, efficacy analyses models will include region as a factor as described in [Section 8](#).

#### **4.5 Multiple comparisons/multiplicity**

The statistical analysis of the primary efficacy variable and selected secondary efficacy variables will account for multiplicity by using a fixed sequence testing procedure. Further details on how the hierarchical testing will be implemented is provided in [Section 8](#).

#### **4.6 Use of an efficacy subset of subjects**

The PPS will be used to evaluate subjects who have efficacy data during the double-blind treatment period and are reasonably compliant with the conditions of the study. This analysis set will provide additional information on the efficacy analysis and will describe findings in a subset of subjects who more closely follow the intentions of the study protocol.

The Randomized Set will be used to evaluate all subjects who were randomized to double-blind study medication. This analysis set will provide additional information on the efficacy analysis by describing findings in the full set of subjects who were randomized and will not exclude those who did not receive at least 1 dose of study medication or did not have a valid Baseline efficacy measurement for ASDAS.

Other than the planned sensitivity analyses based on the PPS and Randomized Set, no other efficacy subsets are defined for statistical analyses during the double-blind 52-week treatment period. The SFE-FAS is defined for analysis of efficacy data at the Week 156/SFE-WD visit.

#### 4.7 Active-control studies intended to show equivalence

Not applicable.

#### 4.8 Examination of subgroups

Subgroups for age (18 – < 45 years,  $\geq 45$  years), gender (male, female), race (white, other), symptom duration (<5,  $\geq 5$  years), tobacco use (never, current former), HLA-B27 genotype (positive, negative), region (North America, Europe, Asia), prior anti-TNF exposure (yes, no), Baseline SPARCC score (<5,  $\geq 5$  as observed in the 2<sup>nd</sup> reading session or in the 1<sup>st</sup> if the 2<sup>nd</sup> reading is not available), subjects with/without relevant changes to background medication, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+) will be utilized for summarizing the primary efficacy variable. Only summary statistics, estimates and confidence intervals will be presented and subjects with missing category will be excluded.

It is possible, that classification errors may occur within the randomization where a subject is found to have been randomized according to the wrong MRI/CRP strata when their MRI data and CRP results are re-examined post randomization. For this reason, the MRI/CRP classification, which will be used in all analyses and subgroup analyses will be one recalculated using the Screening MRI results and second Screening visits CRP results (3 to 5 days prior to Baseline). If the CRP result from the second screening visit is missing, available results from the first screening visit will be used.

Subjects will be categorized as MRI + if they have active (acute) inflammation on MRI highly suggestive of sacroiliitis associated with SpA as determined by the central readers. Otherwise they will be classed as MRI -.

Subjects will be categorized as CRP+ if their second Screening visit CRP result is  $>ULN$ . CRP values  $\leq ULN$  will be considered CRP-. If the CRP result from the second screening visit is missing, available results from the first screening visit will be used. In the unlikely event that a subject has an unavailable recorded CRP result pre Baseline, they will be considered CRP+ for analyses of efficacy but missing for other CRP summaries.

Based on these definitions, the analysis variable for MRI/CRP classification will have the following 3 levels:

- MRI+/CRP+
- MRI+/CRP-
- MRI-/CRP+

Any subjects found to be MRI-/CRP- will have their MRI and CRP data examined to see which of the category thresholds they were close to. The subject will still be included in the FAS analyses however they will be assigned into the respective positive category (MRI+/CRP+, MRI+/CRP- or MRI-/CRP+) that they were closest to and documented to have failed inclusion criterion #9. Hence, they will be excluded from per protocol analyses.

Symptom duration will be categorized into  $<5$  and  $\geq 5$  years for subgroup analyses and calculated as:

Start date of symptoms – date of first study medication administration

The start date of symptoms will be found using the medical history of the subject using the following rules (including imputation of partial dates as described in [Section 4.2.4](#)):

- Subjects with a medical history preferred term of ‘Back pain’ or ‘Inflammatory pain’ will use the earliest start date of these symptoms as the start date of symptoms.
- Otherwise, subjects with a preferred term of ‘Axial spondyloarthritis’ will use the earliest start date of disease as the start date of symptoms
- Otherwise, the subject’s medical history will be reviewed in the DEM meeting to determine if the subject has evidence of a start date of symptoms or primary disease.
- Otherwise, the median duration of subjects with a valid symptom duration within the same stratum regarding MRI/CRP classification, will be used as the symptom duration of the subject, unless the duration is  $<3$  months in which case 3 months and 1 day will be used.

## 5 STUDY POPULATION CHARACTERISTICS

### 5.1 Subject disposition

The number of subjects screened, and reason for screen failure will be summarized for all screened subjects. The disposition of screened subjects will also be presented including the date of the first and last subject into the study by region and site and number of subjects in each population by region and site.

The number of subjects randomized, completed Visit 28/Week 52, entered SFE, completed SFE, prematurely discontinued double-blind study treatment, prematurely discontinued from the study (by phase [main/SFE] and period [DB/OL-CZP/OL-OT]), and reason for premature discontinuation of study treatment and phase (main/SFE) will be summarized by treatment group. The number of subjects who complete each visit (as captured in the eCRF) will also be summarized. In addition, the number of subjects in each of the analysis sets, as specified in [Section 3.5](#), will be summarized as well.

### 5.2 Protocol deviations

The process for reviewing and identifying important protocol deviations is outlined in [Section 3.4](#). The number of subjects with at least one important protocol deviation will be summarized by treatment group in addition to whether that protocol deviation led to exclusion from the PPS.

The subject data listing, in contrast to the general approach (i.e., sorted by treatment group, region, site, subject number), will be sorted by region, site, treatment group, and subject number.

## 6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

The demographics and other Baseline characteristics will be presented for the RS. If the RS differs from the SS or the FAS, then demographics and Baseline characteristics tables may be repeated separately for these analysis sets. These summaries will be presented by treatment group and all subjects.

Tables on medical history and prior diseases, and concomitant and prior medications will also be presented.

## 6.1 Demographics

The following continuous demographic variables will be summarized:

- Age (years)
- Weight (kg)
- Height (cm)
- BMI (kg/m<sup>2</sup>)

The following categorical demographic variables will also be summarized:

- Gender (Female, Male),
- Race (American Indian / Alaskan Native, Asian, Black, Native Hawaiian or Other Pacific Islander, White, Other / Mixed),
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino),
- Racial subgroup (White, Other),
- Age class (to be summarized 3 ways: 1) 18-<65, 65-<85 and >=85; 2) ≤18, >18 to <65 years, and ≥65 years; and 3) 18 – <45 years, 45 – <65 years, ≥65 years),
- BMI class (<18.5 kg/m<sup>2</sup>, 18.5 kg/m<sup>2</sup> to <25 kg/m<sup>2</sup>, 25 kg/m<sup>2</sup> to <30 kg/m<sup>2</sup>, ≥30 kg/m<sup>2</sup>)

## 6.2 Other Baseline characteristics

In this section, the variables that will be summarized as Baseline characteristic are described.

The different binary items of the ASAS criteria will be presented in frequency tables:

- back pain of ≥ 3 month duration and age of onset < 45 (and per inclusion criteria back pain of ≥ 12 month duration and age of onset < 45)
- active (acute) inflammation on MRI highly suggestive of sacroiliitis associated with SpA
- HLA-B27 positivity (current features)
- inflammatory back pain (history and current features)
- arthritis (history and current features)
- enthesitis (history and current features)
- uveitis (history and current features)
- dactylitis (history and current features)
- psoriasis (history and current features)
- Crohn's disease/ ulcerative colitis (history and current features)
- elevated CRP (history and current features)

Symptom duration (calculated as described in [Section 4.8](#)) and time since diagnosis of disease will be summarized as continuous variables and dichotomized into <5 years and  $\geq 5$  years. Time since diagnosis of disease will be defined as: Earliest start date of the medical history (lower level term) of Axial Spondyloarthritis – date of first study medication administration. If a subject does not have a history of ‘Axial Spondyloarthritis’, then the time since diagnosis of disease will be set equal to the symptom duration.

Frequency tables will be provided for the following Baseline variables related to TB status:

- Contact with an individual with active TB
- Contact with an individual who has recently been treated for TB

Summary statistics will be provided for the laboratory variable CRP. Additionally, CRP at Baseline will be categorized based on the number of subjects with CRP  $\leq$ ULN and  $>$ ULN at Baseline.

Frequency tables will summarize HLA-B27, the test at Screening to rule out hepatitis B surface antigen, and the test at Screening to rule out antibodies to hepatitis C.

Frequencies will also be provided for subjects with the following:

- Peripheral arthritis (swollen joint count  $> 0$ ) at Baseline
- Enthesitis (MASES  $> 0$ )
- Extra-articular manifestations:
  - History of uveitis
  - History of psoriasis
  - History of IBD
- Concomitant DMARDs at Baseline
- Prior DMARDs and prior NSAIDs
- MRI/CRP classification (presented as used for stratification at IXRS randomization and as found in the source data along with a cross tabulation to display the number of misclassifications):
  - MRI+/CRP+
  - MRI+/CRP-
  - MRI-/CRP+
- Baseline cardiovascular risk elements (such as smoking, caffeine, and alcohol use)

Note that the term DMARDs is used above as it is common to rheumatic diseases. However, there is currently no conclusive evidence that DMARDs are in fact disease-modifying in axSpA (unlike in RA). As a result, the term SAARDs (slow-acting anti-rheumatic drugs) is used in the protocol to refer to this class of medications. The DMARDs terminology is generally used in this SAP, but it is recognized that SAARDs may be more appropriate in the context of a study of subjects with axSpA.

Since the procedures and surgeries in the procedural history will not be coded, these data will only be presented in listings.

Baseline information for other key measures related to axSpA (eg, BASFI, BASDAI, BASMI, and ASDAS) will be presented in the respective tables providing descriptive statistics over time.

### 6.3 Medical history and concomitant diseases

Previous and ongoing medical history will be summarized by MedDRA<sup>®</sup> system organ class (SOC) and preferred term (PT). Medical procedures are not coded. Concomitant diseases will be recorded by the Investigator at the Screening visit only.

### 6.4 Prior and concomitant medications

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

Concomitant medications will be assigned to treatments according to whether they were taken at least one day in common with the treatment dosing period of that treatment as described in [Table 6–1](#). Dosing periods are defined as follows:

- For double-blind period: From first administration of double-blind medication to last administration of double-blind medication +14 days or first administration of open-label CZP, open-label other treatment or SFE period treatment, whatever comes first.
- For open-label CZP period: From first administration of open-label CZP medication to last administration of open-label CZP medication +14 days or first administration of open-label other treatment or SFE period treatment, whatever comes first.
- For open-label other treatment period: From first administration of open-label other treatment medication to last administration of open-label other treatment medication.
- For SFE period: From first administration of SFE period treatment to last administration of SFE period treatment + 14 days.

A medication can be defined as both prior and concomitant and can be concomitant to more than one treatment. Where a medication start or stop date is (partially) missing, the medication will be considered a concomitant medication, if there is the possibility of concomitant use according to the rules described in [Section 4.2.4](#). Hence, if a partial date implies the concomitant medication could have been taken concomitant to multiple treatments, it will be concomitant to all applicable treatments.

<b>Table 6-1: Assignment of Concomitant Medications to treatments</b>		
Period	Description	Treatment
Double-blind treatment	<ol style="list-style-type: none"> <li>1) If the start or end date of the concomitant medication falls into the double-blind dosing period or</li> <li>2) if the concomitant medication starts prior to double-blind dosing period and ends after the double-blind dosing period or is ongoing,</li> </ol> then assign to the double-blind treatment.	Summarized according to randomized CZP 200mg Q2W or PBO.  For subjects randomized to CZP 200mg Q2W, also include the medications in Total CZP incl DB and OL.  All double-blind period medications will be included in an All subjects summary.
Open-label CZP treatment	<ol style="list-style-type: none"> <li>1) If the start or end date of the concomitant medication falls into the OL-CZP dosing period or</li> <li>2) if the concomitant medication starts prior to the OL-CZP dosing period and ends after the OL-CZP dosing period or is ongoing,</li> </ol> then assign to the OL- CZP treatment.	Summarized according to CZP 200mg Q2W or PBO followed by open-label CZP (i.e. CZP->OL CZP or PBO->OL CZP). All open-label CZP period medications will also be included in a Total CZP incl DB and OL and an All subjects summary.
Open-label other treatment (OL-OT)	<ol style="list-style-type: none"> <li>1) If the start or end date of the concomitant medication falls into the OL-OT dosing period or</li> <li>2) if the concomitant medication starts prior to the OL-OT dosing period and ends after the OL-OT dosing period or is ongoing,</li> </ol> then assign to the OL-OT.	Only included in an All subjects summary.
Safety Follow-up Extension treatment	If the start date of the concomitant medication falls into the SFE dosing period then assign to the SFE treatment.  <b>Note:</b> Concomitant medication information is not systematically recorded during the SFE period. Only concomitant medication taken due to an adverse event will be documented in the CRF during that period. Only new medication starting during the SFE dosing period will therefore be considered.	Not summarized in tables

If a medication starts on the start date of a new dosing period, it will be considered concomitant to the treatment of that new period, but not to the treatment of the previous dosing period (even if the end date of the corresponding dosing period is on the same day as the start of the new dosing period.).

If a medication stops on the start date of a new dosing period it will not be considered concomitant to the treatment of that new dosing period, unless start and stop date of the medication are on the same date.

Past medication summaries will be generated for DMARDs, NSAIDs, and Anti-TNFs by ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and Preferred Term).

Prior DMARDs and NSAIDs will be summarized by WHO-DRL ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and Preferred Term). No extra table for prior Anti-TNFs will be produced due to the fact that this summary is identical to the 1 for the past Anti-TNF.

Prior medication (except DMARDs and NSAIDs) will be summarized by WHO-DRL ATC code (level 2 (3-digit) and level 3 (4-digit) decode).

Concomitant DMARDs and NSAIDs will be summarized by WHO-DRL ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and Preferred Term) for the study as a whole and separately for those medications taken during the SFE for the SFE Safety Set.

Concomitant medication (except DMARDs and NSAIDs) will be summarized by WHO-DRL ATC code (level 2 (3-digit) and level 3 (4-digit) decode).

## 7 MEASUREMENTS OF TREATMENT COMPLIANCE

All described calculations of treatment compliance will correspond to double-blind study medication only.

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 injections. Two syringes will be dispensed in each kit. Both 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. For calculation of total expected syringes, all scheduled administration up to and including the last administration prior to switch to open-label CZP / other treatment or start of SFE treatment will be considered, assuming a 14 days dosing interval. If a subject continues to take double-blind medication after week 52 and prior to start of SFE following instructions from the sites, such administrations will also be included as expected syringes in the calculation.

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 ( $<0.80$ ,  $\geq 0.80 \leq 1.0$ ,  $>1.0 \leq 1.2$  and  $>1.2$ ). The general formula for the compliance ratio (CR) is given as follows:

$$\text{CR for syringes} = \# \text{ actual syringes} / \# \text{ expected syringes}$$

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 ( $<0.80$ ,  $\geq 0.80 \leq 1.0$ ). The general formula for the compliance ratio is given as follows:

$$\text{CR for injection day} = (\text{DB Study Duration} - \text{Cumulative Difference}) / \text{DB Study Duration}$$

The CR for injection day ranges between 0 and 1.

To calculate double-blind study duration, the date of the last injection date prior to double-blind study treatment discontinuation will be compared to the Baseline date, as shown below:

Double-blind Study Duration (days) = last DB injection date – Baseline date (maximum value is 350 days)

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

The sum will be calculated for the 26 visits from Week 0 (Baseline) to Week 50.

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

All efficacy analyses of the double-blind 52-week treatment period will be performed using the FAS. The PPS and Randomized Set will be used for a sensitivity analysis on the primary endpoint only (using the composite endpoint analysis). The SFE-FAS will be used for analysis of efficacy data at the Week 156/SFE-WD visit.

The statistical analysis of the primary efficacy variable and selected secondary efficacy variables will account for multiplicity by using a fixed sequence testing procedure. The predefined order of hypotheses testing, each at a 2-sided 0.05 significance level for the comparison between CZP 200mg Q2W and placebo, will be performed in the sequence shown below. After the first test, each subsequent test is performed only if the previous test is significant in favor of CZP at the 2-sided 0.05 level. Note that the first test relates to the primary efficacy variable while the others provide the sequential testing for selected secondary efficacy variables:

Primary:

1. ASDAS-MI response at Week 52 (Composite endpoint analysis)

Secondary:

2. ASAS40 response at Week 12 (Composite endpoint analysis)
3. Change from Baseline in BASDAI at Week 12 (Reference based multiple imputation analysis)
4. Change from Baseline in BASFI at Week 12 (Reference based multiple imputation analysis)
5. ASAS40 response at Week 52 (Composite endpoint analysis)
6. Change from Baseline in BASDAI at Week 52 (Reference based multiple imputation analysis)
7. Change from Baseline in BASFI at Week 52 (Reference based multiple imputation analysis)
8. Change from Baseline in SI joint SPARCC score at Week 12 (Reference based multiple imputation analysis)
9. Change from Baseline in ASQoL at Week 52 (Reference based multiple imputation analysis)

10. Change from Baseline in nocturnal spinal pain (NRS) at Week 52 (Reference based multiple imputation analysis)

11. Number of subjects with AU or new AU flares through Week 52

Hierarchical testing of efficacy variables for Canada (and any other country where applicable or where requested by Regulatory Authorities):

Primary:

1. ASAS40 response at Week 12 (Composite endpoint analysis)

Secondary:

2. ASDAS-MI response at Week 52 (Composite endpoint analysis)

3. Change from Baseline in BASDAI at Week 12 (Reference based multiple imputation analysis)

4. Change from Baseline in BASFI at Week 12 (Reference based multiple imputation analysis)

5. ASAS40 response at Week 52 (Composite endpoint analysis)

6. Change from Baseline in BASDAI at Week 52 (Reference based multiple imputation analysis)

7. Change from Baseline in BASFI at Week 52 (Reference based multiple imputation analysis)

8. Change from Baseline in SI joint SPARCC score at Week 12 (Reference based multiple imputation analysis)

9. Change from Baseline in ASQoL at Week 52 (Reference based multiple imputation analysis)

10. Change from Baseline in nocturnal spinal pain (NRS) at Week 52 (Reference based multiple imputation analysis)

11. Number of subjects with AU or new AU flares through Week 52

The approach for handling missing values is described in [Section 4.2](#).

The primary analyses to be used in the fixed sequence testing procedure are described in [Sections 8.1](#) and [Section 8.2](#). Further sensitivity analysis and various approaches for handling missing data values are described in [Section 4.2](#).

## **8.1 Statistical analysis of the primary efficacy variable**

The primary efficacy variable measurement for this study is based on ASDAS-MI response at Week 52. However, the primary outcome of the study will be defined as a composite endpoint that is achieved if a subject fulfills the following 2 components:

1. Remain in the study and on the double-blind study treatment through 52 weeks
2. Achieve an ASDAS-MI response at 52 weeks

For simplicity, this primary efficacy variable will be referred to as ASDAS-MI response at Week 52. However, the composite definition as described above will apply when this endpoint is analyzed.

### 8.1.1 Derivation of primary efficacy variable

The ASDAS is comprised 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 listed:

- 0.121 x Back pain (BASDAI Q2 result)
- 0.058 x Duration of morning stiffness (BASDAI Q6 result)
- 0.110 x Patient's Global Assessment of Disease Activity (PGADA)
- 0.073 x Peripheral pain/swelling (BASDAI Q3 result)
- 0.579 x (natural logarithm of the (CRP [mg/L]+ 1))

Back pain, PGADA, duration of morning stiffness, peripheral pain/swelling and fatigue 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 exactly 1 component for the ASDAS is missing at a given visit, that component will be imputed by LOCF, 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-Low Disease (ASDAS-LD): ASDAS  $\geq 1.3$ , <2.1  
\* ASDAS-MD is equivalent to ASDAS-LD. Whereas protocol amendment 5 uses ASDAS-LD throughout the SAP and Shells will continue to use ASDAS-MD in order to maintain consistency with the double-blind TFLs previously produced for this study.
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS  $\geq 2.1$ ,  $\leq 3.5$
- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

Note that ASDAS-MD (used in study protocol amendments 4 and earlier) is identical to the definition of ASDAS-LD used in the study protocol amendment 5. ASDAS-LD instead of ASDAS-MD is following an update of the nomenclature (Machado et al, 2018). ASDAS-MD is used throughout this SAP and in the analysis outputs in order to retain consistency with the Week 52 interim database lock analyses.

The variables related to ASDAS improvement are defined as follows:

- ASDAS-Clinically Important Improvement (ASDAS-CII): ASDAS reduction (improvement) of  $\geq 1.1$  relative to Baseline
- ASDAS-Major Improvement (ASDAS-MI): ASDAS reduction (improvement) of  $\geq 2.0$  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).  
Note: As a sensitivity analysis, the primary analysis will be repeated where ASDAS-MI includes the criteria ASDAS reduction (improvement) of  $\geq 2.0$  relative to Baseline only.

ASDAS-MI at 52 weeks is the primary variable for efficacy analysis.

### 8.1.2 Primary analysis of the primary efficacy variable

The primary analysis for of the primary efficacy variable will be based on logistic regression. The odds ratio of the ASDAS-MI responder rates at Week 52 will be estimated and tested between treatment groups using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+). The MRI/CRP classification used in all statistical analysis will be recalculated using source data as described in [Section 4.8](#) and the IXRS MRI/CRP categorization will not be used. If the logistic regression model is unable to converge, then the MRI/CRP classification variable may be dropped from the model to facilitate convergence. The study design is likely to result in some sites with a small number of enrolled subjects. The use of region as a factor in the model is intended to combine study centers in similar geographic regions. The geographic regions to be used will be as defined in the IXRS randomization system. In the unlikely event that a subject moves sites during the study which leads to a change in region, the region originally used at randomization will be used in the analysis.

Given the composite endpoint definition described in [Section 8.1](#), there will be no missing data for the primary endpoint, as subjects that discontinue double-blind study treatment prior to Week 52 or who do not have an ASDAS-MI response at Week 52 are considered non-responders to the double-blind study treatment.

Tables will present the responder rates for placebo and CZP 200mg, the respective effect estimates (adjusted odds ratio with reference to placebo), p-values, and 95% confidence intervals (CIs). The odds ratios and 95% CIs will also be presented graphically.

### 8.1.3 Supportive and sensitivity analyses of the primary efficacy variable

As described in [Section 6.3](#) of the protocol, subjects who discontinue study treatment will not necessarily be withdrawn from the study. In an attempt to minimize missing data, efforts will be made to continue to collect safety and efficacy data on these subjects through study completion. In order to assess the impact of various missing data assumptions on the analysis, four sensitivity analyses of the primary efficacy variable will be performed as described in [Section 4.2.1](#). Further details of these analyses are provided in this section. The odds ratios  $\pm$  95% CI's for the primary analyses and sensitivity analyses will be presented on the same graph to allow visual comparison of the robustness of the results. In addition, a summary of the missing data pattern will be provided.

For Canada (and any other country where applicable or where requested by Regulatory Authorities), the primary efficacy variable measurement is the ASAS40 response at Week 12. Analysis of the ASAS40 response at Week 12 will be performed as described above for

ASDAS-MI response at Week 52 (including all sensitivity and subgroup analyses). Similar to the ASADAS-MI efficacy endpoint, the primary analysis of ASAS40 will be treated as a composite endpoint when it is analyzed. Specifically, subjects will need to: 1) achieve the relevant response (ASAS40); and 2) remain in the study and on their randomized double-blind study treatment through Week 12.

In addition to the above analyses, subgroup summaries using descriptive statistics, estimates and confidence intervals only for the primary efficacy variable by age, gender, race, symptom duration, tobacco use, HLA-B27 genotype, region, prior anti-TNF exposure, Baseline SPARCC score ( $<5$ ,  $\geq 5$ ), subjects with/without relevant changes to background medication, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+) will be performed.

The logistic regression analyses of the composite endpoint for ASDAS-MI responder rates at Week 52 (as described in [Section 8.1.2](#)) will also be repeated using the PPS and the Randomized Set. In addition, the primary analysis will be repeated where ASDAS-MI includes the criteria ASDAS reduction (improvement) of  $\geq 2.0$  relative to Baseline only.

### 8.1.3.1 Including observed data at Week 52

The observed data at Week 52 will be used and analyzed according to randomized treatment, regardless of whether or not the subject is still on his/her randomized double-blind study treatment at that time. For example, subjects on OL-CZP or OL-OT at the time of their Week 52 visit will have that data included in the analysis including observed data at Week 52. Despite efforts to continue to collect data on all subjects (even if they discontinue the double-blind study treatment), there may still be subjects with missing data at Week 52. Such subjects will be considered as non-responders. The same logistic regression model specified for the primary analysis will be used. The odds ratio of the treatment comparison and corresponding 95% CI for ASDAS-MI responder rates at Week 52 based on the observed data will be estimated using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+).

### 8.1.3.2 Multiple imputation

Investigators will be given discretion to discontinue study treatment for subjects not responding in order to provide them with other treatments for the condition. Therefore, in many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent with a missing at random (MAR) pattern of missingness. In order to investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied as follows:

1. The continuous ASDAS score at each visit will be calculated using the methods described in [Section 8.1.1](#) (including the application of LOCF if only one component is missing). Visits with more than one component missing will be set to missing. ASDAS data collected after the discontinuation of double-blind study treatment will be treated as missing in the original data set for this sensitivity analysis. Hence, all open-label period data is excluded from this analysis.
2. For subjects with a non-monotone (intermittent) missing pattern, MCMC will be used to impute their missing intermittent ASDAS scores including treatment arm, region, MRI/CRP

classification and ASDAS measurement from previous visit(s) as explanatory variables in the model. This will be done 100 times limiting the imputation to in range results.

After this step there will be 100 datasets containing all subjects however the only missing data will be a monotone missing data pattern. A variable called “\_imputation\_” will identify each of the 100 imputed datasets output.

3. The remaining missing data for all subjects with a monotone missing data pattern will be imputed using regression with treatment arm, region, MRI/CRP classification and ASDAS measurement from previous visit(s) as explanatory variables. This model will be run once for each of the imputed datasets limiting imputation to in range results.
4. The ASDAS change from Baseline to Week 52 will be calculated using the complete datasets (with no missing data) and categorized as: ASDAS-MI or not ASDAS-MI (as defined in [Section 8.1.1](#))
5. For each of the 100 imputed datasets (identified by \_imputation\_ variable), the odds ratio of the treatment comparison and corresponding 95% CI for ASDAS-MI responder rates at Week 52 based on the observed data will be estimated using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+) and Baseline as a covariate in the same manner as the primary endpoint analysis.
6. The results from these analyses will be combined into a single inference using SAS PROC MIANALYZE.

The parameter estimate and 95% CIs will be back transformed (using the exponential function) to give the odds ratio and CI for the CZP vs PBO comparison. This will be presented alongside the p-value from the proc mianalyze.

### 8.1.3.3 Tipping point analysis

In this analysis, various assumptions will be made about average outcomes among the subsets of subjects who prematurely discontinued double-blind study treatment and hence have a monotone missing data pattern. Various “delta adjustments” will be made to the assumed responses among missing ASDAS data in each treatment arm with varying degrees of plausibility (O’Kelly, 2014) prior to classification of the composite endpoint ASDAS-MI. The goal of the tipping point analysis is to identify assumptions about the missing data under which the conclusions from the logistical regression change, i.e., under which there is no longer evidence of a treatment effect. Then, the plausibility of such assumptions will be discussed.

Since the tipping point analysis is to find out how much of a change is required to “tip” the result from statistically significant at the  $\alpha=0.05$  level to not statistically significant, it will not be performed if the primary endpoint analysis is not statistically significant ( $p\text{-value} \leq 0.05$ ) at the onset.

1. Steps 1 and 2 from [Section 8.1.3.2](#) will be completed followed by steps 2-6 as shown below. Note that as described in [Section 8.1.3.2](#), part 1, ASDAS data collected after the discontinuation of double-blind study treatment will be treated as missing and then imputed using the methods described. Hence, all open-label period data is excluded from this analysis.

2. The remaining monotone missing data for all subjects will be imputed by treatment group. A regression with ASDAS measurement from the previous visit(s), MRI/CRP classification and region as explanatory variables will be used to impute the missing data including a MNAR mechanism under different assumptions. For subjects randomized to the Placebo treatment group, a MNAR mechanism will be assumed by using a specified negative shift to decrease the ASDAS parameter values prior to classifying the subject according to the ASDAS-MI endpoint. For subjects randomized to the CZP treatment group, a MNAR mechanism will be assumed by using specified positive shift parameters to increase imputed ASDAS values prior to classifying the subjects according to the ASDAS-MI endpoint. The size of the shift in the CZP or PBO treatment group will be investigated to determine at what point the analysis leads to a non-significant result.

As it is the raw ASDAS scores being imputed, the shift size (delta) will start with 0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6 and 1.8. Further values for delta may be investigated depending on the results obtained in order to determine the point at which the results are no longer significant.

When tipping point analyses is applied to ASAS40 response, the imputation will be done on the raw components before assessing the imputed results for ASAS40 response. The same delta shifts (0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6 and 1.8) will be investigated applied to the raw components (each of which are on a scale from 0 to 10) prior to calculating ASAS40 response. Further values for delta may be investigated depending on the results obtained.

This will be done once for each of the 100 datasets. The same delta value will be assumed for all visits.

3. The ASDAS change from Baseline to Week 52 will be calculated using the complete datasets (with no missing data) and categorized as: ASDAS- MI or not ASDAS-MI.
4. For each of the 100 imputed datasets (identified by `_imputation_` variable), the odds ratio of the treatment comparison and corresponding 95% CI for ASDAS-MI responder rates at Week 52 based on the observed data will be estimated using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+) and Baseline as a covariate in the same manner as the primary endpoint analysis.
5. The results from these analyses will be combined into a single inference using SAS PROC MIANALYZE. The parameter estimate and 95% CIs will be back transformed (using the exponential function) to give the odds ratio and CI for the CZP vs PBO comparison. This will be presented alongside the p-value from the proc mianalyze.
6. Steps 2 through 5 will be repeated using different values for the shift parameter as specified in step 2. Following the final Step 5, for each value of delta, the odds ratios, CIs and p-values will be presented in a single table with the associated delta value which was applied in the analysis.

#### 8.1.3.4 Observed case analysis

This analysis will only include the observed data for subjects still on the original double-blind study treatment. Subjects with missing ASDAS data at Week 52 will not be included in the

analysis and hence only a subset of subjects will be analyzed and no composite endpoint implemented. The same logistic regression model specified for the primary efficacy analysis will be performed. The odds ratio and corresponding 95% CI of the ASDAS-MI responder rates at Week 52 based on the observed data will be estimated and tested between treatment groups using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+).

## 8.2 Statistical analysis of secondary efficacy variables

The secondary efficacy variables are designated as variables 2 through 11 in the hierarchical testing procedure outlined in [Section 8](#), in addition to counts of relevant changes to background medication. The statistical methodology to be used for the analysis of these variables is described below.

### 8.2.1 Derivation of secondary efficacy variables

#### 8.2.1.1 Assessment in Axial SpondyloArthritis international Society response criteria (ASAS20/40, ASAS5/6, and ASAS partial remission)

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 BASFI);
- Inflammation (the mean of the BASDAI questions [Q] 5 and 6) concerning morning stiffness intensity and duration)

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].

ASAS data collected after the discontinuation of double-blind study treatment will be treated as missing in the original data set.

The ASAS criteria for 40% improvement (ie, ASAS40) are defined as relative improvements 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 5/6 response is defined as at least 20% improvement in 5 of 6 domains, which includes spinal mobility (i.e., lateral spinal flexion of the BASMI) and CRP as more objective measures (Brandt et al, 2004). For CRP values below the lower limit of quantification (<4mg/L), half the lower limit (2 mg/L) will be used as the imputed value.

The ASAS partial remission response is defined as a score of  $\leq 2$  units on a 0 to 10 unit scale in all 4 domains.

ASAS40 response at Week 12 and Week 52 are secondary efficacy endpoints.

#### 8.2.1.2 Bath Ankylosing Spondylitis Disease 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 NRSs, each with 10 units to measure the severity of the 5 major symptoms: fatigue, spinal pain, peripheral joint pain and swelling, enthesitis, and morning stiffness (both severity and duration) over the last week. To give each symptom equal weighting, the average of the 2 scores relating to morning stiffness 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.

If 1 of the 2 morning stiffness measurements (ie, questions: “How would you describe the overall level of morning stiffness you have had from the time you wake up?” and “How long does your morning stiffness last from the time you wake up?”) is missing, the other one will be used for the morning stiffness calculation. The same imputation is also applied for the calculation of the ASAS inflammation component, which is calculated as the average of the 2 morning stiffness 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.

The BASDAI data collected after the discontinuation of double-blind study treatment will be treated as missing in the original data set.

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

Change from Baseline in BASDAI at Week 12 and Week 52 are secondary efficacy endpoints.

#### **8.2.1.3 Bath Ankylosing Spondylitis Functional Index (BASFI)**

The BASFI contains 10 questions. The first 8 questions evaluate activities related to functional anatomical limitations due to the course of this inflammatory disease. The final 2 questions evaluate the subjects’ ability to cope with everyday life. An NRS ranging from 0 to 10 is used to answer the questions on the test.

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

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.

The BASFI data collected after the discontinuation of double-blind study treatment will be treated as missing in the original data set.

Change from Baseline in BASFI at Week 12 and Week 52 are secondary efficacy endpoints.

#### **8.2.1.4 Magnetic Resonance Imaging (MRI) assessments**

The SPARCC scoring method for lesions found on the MRI is based on an abnormal increased signal on the short-tau-inversion recovery (STIR) sequence, representing bone marrow edema (defined as an increased signal in bone marrow on a T2-weighted sequence, reflecting an increased concentration of “free water” related to a bone lesion). Each SI joint is divided into 4 quadrants: upper iliac, lower iliac, upper sacral, and lower sacral. The presence of increased signal on STIR in each of these 4 quadrants is scored on a dichotomous basis, where 1 =

increased signal and 0 = normal signal. Joints that include a lesion exhibiting intense signal are each given an additional score of 1 per slice that demonstrated this feature. Similarly, each joint that included a lesion demonstrating continuous increased signal of depth greater or equal 1 cm from the articular surface is also given an additional score of 1. The scoring is repeated in each of 6 consecutive coronal slices. Total SI joint SPARCC scores can range from 0 to 72.

The Berlin modification of the ASspiMRI-a is a scoring system with a concentration on STIR sequences without other fat saturation techniques. This scoring method quantifies active changes in 23 vertebral units (VU) of the spine (from C2 to S1). A VU is defined as the region between 2 virtual lines through the middle of each vertebra. Active inflammation is scored by grading the degree of bone marrow edema from 0 to 3 in 1 dimension on 1 or more consecutive slices that represent the highest level of inflammation in a particular VU. Total spine ASspiMRI-a score in the Berlin modification can range from 0 to 69.

The following imputation rules should be used for calculating the total for both SI joint SPARCC scores and spine ASspiMRI-a score in the Berlin modification:

- If all scores are NA at a visit, the imputed total is blank for that visit.
- Treat NA as 0 when computing the total score.
- Carry the NA score from the Baseline visit forward to all post-Baseline visits unless all scores at Baseline are NA.
- Carry the numeric score from the last visit with non-NA score forward if a score is NA at a post-Baseline visit (unless all scores are NA post-Baseline).
- If ALL the Baseline scores are NA, then do not carry forward the Baseline scores. Treat the subsequent visit as a surrogate Baseline.

The details related to how the MRI data will be read and adjudicated will be outlined in a separate imaging charter. In the cases where the initial two readings are discrepant, a third reading is performed by an independent reader. This third score is not an adjudication but rather an additional score and value for analysis which will be based on the average of all 3 reader assessments. There will also be two sessions of readings. The first session will conduct readings of the Baseline and Week 12 MRI and will be used for summaries at Week 12. The second session will conduct readings of the Baseline, Week 12 and Week 52 MRIs and will be used for summaries at Week 52.

The following visit windows will be applied:

- All Baseline values should be collected prior to randomization as MRI is required to be randomized into the study.
- With the exception of Week 156, for post-Baseline visits, a time window of 2 weeks before or after the scheduled visit will be used for mapping MRI data to the given visit.
- For Week 156, a time window of 4 weeks before or after the scheduled visit will be used for MRI data to the given visit. The expected day for Week 156 will be calculated assuming that the SFE period start was on Week 52. I.e. Week 156 will be expected to be  $(156-52) \times 7$  days after Week 52. For patients where the first dose after the Week 50 is before the SFE informed consent the width of the lower boundary of the Week 156 window will be increased further

by the absolute number of days between the first dose after the Week 50 visit and the SFE informed consent date.

The SI joint SPARCC data collected after the discontinuation of double-blind study treatment will be treated as missing for any observed case analysis.

The change from Baseline in SI joint SPARCC score at Week 12 is a secondary efficacy endpoint.

#### **8.2.1.5 Relevant changes to background medication**

For purposes of this secondary efficacy variable, relevant changes to background medication will be defined as the following:

- The addition of a new DMARD or the change from one DMARD to another
- The addition of an NSAID or the change from one NSAID to another
- Increased dose of chronic oral corticosteroids
- Increased dose in chronic analgesic medications or the addition of a chronic analgesic medication

The variable to be measured is the number of subjects without relevant changes to background medications. Therefore, a subject who does not have any of the above relevant changes made to background medication during the study through Week 52 will be considered to have met this endpoint. Conversely, subjects who have one of the above relevant background medication changes or who do not complete double-blind study treatment to Week 52 will be considered as not having met this endpoint.

#### **8.2.1.6 Ankylosing Spondylitis Quality of Life (ASQoL)**

The ASQoL, a validated disease-specific 18-item questionnaire, has been developed specifically for measuring HRQoL in subjects with AS (Doward et al, 2003). A nonradiographic axSpA (nr-axSpA) specific scoring algorithm is presently being developed, but it is unclear when this algorithm will become available relative to the timing of the AS0006 database lock. Should this alternative scoring approach become available prior to finalization of the clinical study report, analyses of ASQoL data may be repeated using this alternate scoring approach appropriate for nr-axSpA population and methods will be described either in a SAP amendment or within the Documentation of Statistical Methods appendix to the study report.

#### **8.2.1.7 Nocturnal Spinal Pain**

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.

#### **8.2.2 Primary analysis of the secondary efficacy variables**

The comparison of CZP 200mg Q2W to placebo as described in this section for the ASAS40 response, BASDAI, and BASFI at Week 12 and Week 52, SI joint SPARCC score at Week 12, ASQoL at Week 52 and Nocturnal Spinal Pain at Week 52 will be part of the fixed sequence testing procedure outlined in [Section 8](#).

### **8.2.2.1 Primary analysis of categorical secondary efficacy variables**

The ASAS40 response at Week 12 and Week 52 are responder secondary efficacy variables. For Canada and any other country where applicable or where requested by Regulatory Authorities, the ASAS40 response at Week 12 is the primary efficacy variable; the ASAS40 response at Week 52 and the ASDAS-MI response at Week 52 are secondary efficacy variables. Hence the primary and sensitivity efficacy analysis on ASDAS-MI at Week 52 will be repeated for ASAS40 at Week 12.

As a responder variable, ASAS40 will be analyzed using logistic regression in the same manner as for the primary analysis described in [Section 8.1.2](#). As with the primary efficacy variable, the ASAS40 will be treated as a composite endpoint when it is analyzed. Specifically, subjects will need to 1) achieve the relevant response (ASAS40) and 2) remain in the study and on their randomized double-blind study treatment through the time point being analyzed (Week 12 or Week 52).

Relevant changes to background medication will also be analyzed using the same composite endpoint logistic regression model described for the primary variable described in [Section 8.1.2](#).

The number of subjects with new onset post-Baseline AU or new AU flares through Week 52 is a categorical secondary efficacy endpoint for all regions. Subjects will be classified as having post-Baseline Uveitis: a TEAE of Uveitis, a new diagnosis of Uveitis or a flare of uveitis at any visit post-Baseline which is on or before the Week 52 visit. All subjects without evidence of uveitis events, new diagnosis or flares will be considered not to have post-Baseline Uveitis. As a dichotomous outcome variable, treatment groups will be compared for differences using logistic regression including treatment, region and MRI/CRP classification as explanatory variables. Subjects with no evidence of new AU flares will be considered as having no flare.

### **8.2.2.2 Primary analysis of continuous secondary efficacy variables**

The continuous BASDAI, BASFI, ASQoL and Nocturnal Spinal Pain score at each visit will be calculated using the methods described in [Sections 8.2.1.2, 8.2.1.3, 8.2.1.6 and 8.2.1.7](#). Missing data (including data excluded due to subjects discontinuing double-blind study treatment) at each visit will be handled via reference based multiple imputation as described below:

The reference-based multiple imputation assumes that the statistical behavior of the CZP and placebo-treated subjects after discontinuing study medication becomes that of the placebo-treated subjects. Data collected after discontinuation of the double-blind study treatment for both the CZP and placebo groups will be considered missing. Multiple imputations are used to replace missing outcomes for drug- and placebo-treated subjects who discontinued using multiple draws from the posterior predictive distribution estimated from the placebo arm. For binary efficacy variables (ASAS40 at Weeks 12 and 52), imputation will be done on the raw components before assessing the imputed results for ASAS response. For continuous efficacy variables (change from Baseline in BASDAI and BASFI at Weeks 12 and 52, change from Baseline in the SI joint SPARCC score at Week 12, change from baseline in ASQoL and Nocturnal Spinal Pain at Week 52), imputation will be done on the raw scores prior to calculating the change from Baseline.

1. For subjects with a non-monotone (intermittent) missing pattern, MCMC will be used to impute their missing intermittent scores including treatment, region and MRI/CRP

classification as explanatory variables. This will be done 100 times in order to provide a dataset with monotone missing limiting imputation to in range results.

After this step there will be 100 datasets containing all subjects however the only missing data will be a monotone missing data pattern. A variable called “\_imputation\_” will identify each of the 100 imputed datasets output.

2. The remaining missing data for all subjects with a monotone missing data pattern will be imputed using regression using the placebo-treated subjects arm only to impute missing data for both placebo and CZP treatments. The model will contain region, MRI/CRP classification and measurement from previous visit(s) as explanatory variables. This model will be run once for each of the imputed datasets limiting imputation to in range results.
3. For each completed data set, the outcome (response or change from Baseline) will be calculated using the complete datasets (with no missing data).
4. For ASAS40, complete parts 5 and 6 from [Section 8.1.3.2](#) in order to provide odds ratios, 95% CIs and the p-value from the proc mianalyze. For BASDAI, BASFI, SI joint SPARCC, ASQoL and Nocturnal Spinal Pain, complete parts 4 and 5 from [Section 8.2.2.2](#) in order to provide treatment differences, 95% CIs and the p-value from the proc mianalyze.

Additionally, change from Baseline in SI joint SPARCC score at Week 12 is a continuous secondary efficacy variable. The same referenced based multiple imputation procedure specified in this section will be used to account for missing SPARCC data. However, it should be noted that, unlike BASDAI and BASFI, SPARCC score will not be measured at any time points between Baseline and Week 12, meaning that those are the only 2 time points that can be considered in the evaluation of SPARCC score at Week 12. Comparisons between treatment groups will be made using an analysis of covariance (ANCOVA) model on the imputed data set. The model will include Baseline score, treatment group, region, and MRI/CRP classification.

### **8.2.3 Supportive and sensitivity analyses of the secondary efficacy variables**

Descriptive statistics (number of available observations [n], mean, median, standard deviation, minimum and maximum) will be provided for all secondary efficacy variables. The descriptive analyses will be covered in the tables summarizing the variables over time.

Sensitivity analyses of the secondary binary efficacy variables (ASAS40 at Weeks 12 and 52) as described in [Section 8.2](#) will be as follows:

- Including observed data at Week 52 (or Week 12 as applicable): As described for the primary efficacy variable in [Section 8.1.3.1](#).
- Multiple imputation: As described for the primary efficacy variable in [Section 8.1.3.2](#). Note that as ASAS40 is a composite of 4 different variables, the multiple imputation procedure will be performed on each of these components, and the ASAS40 response will be derived prior to Step 5 based on the multiply imputed datasets.
- Reference-based multiple imputation: As described in [Section 8.2.2.2](#).
- Observed case analysis: As described for the primary efficacy variable in [Section 8.1.3.4](#).

The sensitivity analyses of secondary continuous efficacy variables (the change from Baseline in BASDAI and BASFI at Weeks 12 and 52, the change from Baseline in the SI joint SPARCC

score at Week 12, the change from baseline for ASQoL at Week 52 and the change from baseline for Nocturnal Spinal Pain at Week 52) will be as follows:

- Including observed data at Week 52 (or Week 12 as applicable): As described for the primary efficacy variable in [Section 8.1.3.1](#)
- Multiple imputation: As described in [Section 8.2.3.1](#)
- Observed case analysis: As described for the primary efficacy variable in [Section 8.1.3.4](#).
- Last Observation Carried Forward (LOCF): If there is missing data at the time point of interest, the last available post-Baseline measurement will be carried forward to that time point (including scheduled or unscheduled visits). This will not apply for the analysis of SI joint SPARCC score at Week 12 since there are no measurements collected between Baseline and Week 12.

### 8.2.3.1 Multiple imputation of secondary continuous efficacy variables

The continuous BASDAI, BASFI, ASQoL or Nocturnal Spinal Pain scores at each visit will be calculated using the methods described in [Section 8.2.1.2](#), [Section 8.2.1.3](#), [8.2.1.6](#) and [8.2.1.7](#). Missing data (including data excluded due to subjects discontinuing double-blind study treatment) at each visit will be handled via multiple imputation as described below:

1. For subjects with a non-monotone (intermittent) missing pattern, MCMC will be used to impute their missing intermittent scores. This will be done 100 times for each subject including treatment arm, region, MRI/CRP classification and limiting imputation to in range results.

After this step there will be 100 datasets containing all subjects however the only missing data will be a monotone missing data pattern. A variable called “\_imputation\_” will identify each of the 100 imputed datasets.

2. The remaining missing data for all subjects with a monotone missing data pattern will be imputed using regression with treatment arm, region, MRI/CRP classification and measurement from previous visit(s) as explanatory variables. This will be done once for each of the 100 datasets limiting imputation to in range results.
3. The change from Baseline to Week 52 will be calculated using the complete datasets (with no missing data).
4. For each of the 100 imputed datasets (identified by \_imputation\_ variable), an ANCOVA model, where response is the change from Baseline, with Baseline score as a fixed-effect covariate and treatment group, region, and MRI/CRP classification as fixed-effect categorical factors. The estimand of interest here is the difference in BASDAI/BASFI/ASQoL/Nocturnal Spinal Pain, if all subjects tolerate or adhere to double-blind study treatment. In the context of handling missing data in clinical study settings, this has been referred to as “estimand 3” (Mallinckrodt, 2012).
5. The results from these analyses will be combined into a single inference using SAS PROC MIANALYZE.

### 8.3 Analysis of other efficacy variables during the double-blind 52-week treatment period

The efficacy analyses described in this section are not part of the multiplicity-controlled testing procedure described in [Section 8](#). The p-values reported for these analyses will not be adjusted for multiplicity and will be considered nominal. Derivations for these other efficacy variables can be found in [Section 8.3.1](#).

Continuous other efficacy variables will be analyzed using an ANCOVA model including Baseline score as a covariate, and fixed effects for treatment group, region, and MRI/CRP classification. The treatment differences and corresponding 95% CIs will be calculated based upon the adjusted least squares means. Missing values or values observed after discontinuing the double-blind study treatment will be imputed using LOCF. In addition, as described in [Section 4.2.3](#), an observed case analysis will also be performed for subjects still on the original double-blind study treatment. Tables for continuous variables will display descriptive statistics for Baseline, absolute values and change from Baseline at post-Baseline time points.

Categorical other efficacy variables will be analyzed using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+). If the logistic regression model is unable to converge, then the MRI/CRP classification variable may be dropped from the model to facilitate convergence. Tables for the binary variables will display the responder rates for the various time points. Analyses will be presented in 2 ways (as described in [Section 4.2.3](#)). First, they will be presented where the denominator is all subjects in the FAS for the given treatment group. This is essentially a NRI approach as subjects who have not achieved the given outcome (whether observed or not) or subjects who discontinue double-blind study treatment are considered as not having responded. Second, the responder rates will be presented for the FAS using Observed Case data. Subjects still on the original double-blind study treatment and with response will be analyzed, with the denominator based on the number of observed values at the given time point.

Continuous other efficacy variables to be analyzed using change from Baseline are:

- BASDAI and BASFI at Weeks 1, 2, 4, 8, 16, 20, 24, 28, 32, 36, 40, 44, 48
- ASDAS at Weeks 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- ASQoL at Weeks 1, 2, 4, 12, 24, 36, 48
- Nocturnal spinal pain (NRS) at 1, 2, 4, 8, 16, 20, 24, 28, 32, 36, 40, 44, 48
- Other continuous data from the ASAS components at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52:
  - PGADA
  - Total spinal pain (NRS),
  - Morning stiffness as assessed by the average of BASDAI questions 5 and 6
  - BASMI Lateral Spine Flexion
  - CRP

- BASDAI individual questions 1 (Fatigue), 2 (neck, back or hip pain), 3 (pain/swelling in joints other than neck) and 4 (discomfort) at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- BASMI linear score at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- Spinal mobility assessments (occiput to wall distance and chest expansion) at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- Sacroiliitis grading for structural damage at Week 52 (mean of the right and left grade for all readers). Note: LOCF will not be performed as grading is only measured as Baseline and Week 52
- SI joint SPARCC score at Week 52 and ankylosing spine MRI acuity (ASspiMRI-a) in the Berlin modification at Weeks 12 and 52
- ASAS-NSAID use at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52  
Thereby, ASAS-NSAID use at a certain week refers to any intake between this week and the previous week.
- Sleep Problems Index II domains of the MOS Sleep scale at Weeks 4, 12, 24, 36 and 48
- Enthesitis (MASES) at Weeks 1, 2, 4, 8, 12, 16, 24, 32, 40, 48, and 52
- Swollen and tender joint counts at Weeks 4, 12, 24, 36 and 52
- PhGADA at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- SF-36, PCS, MCS and individual domains at Weeks 4, 12, 24, 36, 48, and 52
- EQ-5D VAS at Weeks 4, 12, 24, 36, 48, and 52

Categorical other efficacy variables to be analyzed are:

- ASDAS-MI at Weeks 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48,
- ASDAS-CII at Weeks 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- ASDAS ID, MD, HD and vHD at Baseline and Weeks 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52 (Summary statistics provided only)
- BASDAI50 response at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- ASAS40 at Weeks 1, 2, 4, 8, 16, 20, 24, 28, 32, 36, 40, 44, 48
- ASAS20, ASAS5/6 and ASAS partial remission response at Weeks 1, 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52
- Sacroiliitis grading for structural damage at Week 52 (change from mNY- negative to mNY-positive)
- SI joint SPARCC score <2 at Weeks 12 and 52
- Uveitis flares, IBD exacerbations and Psoriasis exacerbations at Baseline and Weeks 4, 12, 24, 36, 48 and 52 (Summary statistics provided only)
- EQ-5D 5-Item health status at Weeks 4, 12, 24, 36, 48, and 52

For the ‘resource utilization’ variables ‘number of concomitant medical procedures’, ‘number of health care provider consultations not foreseen by the protocol’, ‘number of hospitalizations’, and ‘number of emergency room visits’ descriptive statistics and frequency distributions will be presented for the entire double-blind period.

For the WPS scores, Question 1 responses will be summarized descriptively. Question 2-9 will be summarized by descriptive statistics as well. Treatment comparisons for Question 2 through Question 9 will be performed using a mixed-effect repeated-measures model (MMRM). For each individual question, a separate MMRM will be fitted for the change from baseline, including fixed effects of baseline, treatment, visit, treatment-by-visit interaction and baseline-by-visit interaction, including a random effect of subject and using an unstructured variance-covariance structure. The Kenward-Rogers method will be applied to estimate the denominator degrees of freedom. If convergence problems arise, the covariance structure will be assumed to be autoregressive first order (AR[1]). If that still leads to numerical problems, the more general case, the Toeplitz (TOEP) covariance structure will be used. If all methods above fail to produce a solution, the covariance structure Compound Symmetry (CS) will be used. The least-square means of the treatment difference by visit including 95% confidence intervals and p-values for the test on treatment differences will be summarized.

### 8.3.1 Derivation of other efficacy variables

Derivations of other efficacy variables not covered in the primary or secondary variables are provided below.

#### 8.3.1.1 Bath Ankylosing Spondylitis Metrology Index (BASMI)

The BASMI characterizes the spinal mobility of a subject with AS and consists of 5 clinical measures to reflect axial status: CCI

Each of the 5 movements is scored according the linear BASMI definition (see table below). The mean of the 5 scores provides the BASMI score. The higher the BASMI score the more severe the subject’s limitation of movement due to their AxSpA.

##### BASMI linear definition

$S = (21.1 \text{ cm} - A) / 2.1 \text{ cm}$	CCI
$S = (A - 8 \text{ cm}) / 3 \text{ cm}$	
$S = (7.4 \text{ cm} - A) / 0.7 \text{ cm}$	
$S = (124.5 \text{ cm} - A) / 10 \text{ cm}$	
$S = (89.3^\circ - A) / 8.5^\circ$	
Always with the additional condition $0 \leq S \leq 10$	

S = score, A = assessment

For CCI, take the mean of the left and right measurements, if both are available. Otherwise, the available measurement will be used.

For the CCI, values greater than 9.0 cm (Maksymowych 2006) will be flagged as invalid and treated as if they were missing values. The below imputation rules apply for BASMI.

If 1 or 2 clinical measures for the BASMI are missing at one visit, the missing measure will be imputed by carrying the last observation forward, and the BASMI will be calculated accordingly. If more than 2 items are missing, the BASMI score will be treated as missing.

### **8.3.1.2 Total spinal pain**

Total 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.

### **8.3.1.3 SI joint radiographs**

SI joint radiographs will be read by two reviewers, with adjudication (if applicable), to assess for the presence and severity of sacroiliitis according to mNY criteria for AS in the right and left SI joint separately.

A grading from 0 to 4 will be assigned corresponding to 0 = Normal, 1 = Suspicious but not definite, 2 = Minimal: some sclerosis, minimal erosion, no marked joint space narrowing, 3 = Moderate: definite sclerosis, both sides of the joint with erosions and/or joint space change, and 4 = Ankylosis: complete obliteration of the SI joint with or without sclerosis. Readers will also classify the subjects as mNY-negative or mNY-positive.

A mean grading will be calculated per subject and visit using the four SI grading results (the left and right joint grades read by the 2 readers). If a third adjudication reading is required, then all 6 grades will be used to calculate the mean grade per visit.

If there is disagreement between the 2 readers in the classification of the subject (mNY-negative or mNY-positive), then the 3rd readers adjudication classification will be used for analysis.

Subjects who are positive at baseline (and hence should not have been enrolled) will be excluded from the categorical endpoint analysis of subjects who change from mNY-negative to mNY-positive during the study.

### **8.3.1.4 Tender and Swollen Joint Counts (44 joints evaluation)**

Tender joint counts (TJC) and swollen joint counts (SJC) will be carried out on the following 44 joints:

- 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:

0 = None (not tender)

1 = Positive (tender)

Each joint is scored for swelling as follows:

0 = None

1 = Detectable

TJC and 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.

### 8.3.1.5 Patient's Global Assessment of Disease Activity (PGADA)

PGADA will be recorded based on an NRS describing how active the subject's spondylitis was on average over the past week. The range of the scale is from 0 to 10, where 0 represents not active and 10 represents very active.

### 8.3.1.6 Physician's Global Assessment of Disease Activity (PhGADA)

PhGADA is recorded by the physician on a VAS ranging from 0 to 100, where 0 is "very good, asymptomatic and no limitation of normal activities" and 100 is "very poor, very severe symptoms which are intolerable and inability to carry out all normal activities."

### 8.3.1.7 C-reactive protein (CRP)

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

### 8.3.1.8 Spinal mobility

The following spinal mobility assessments will be performed in addition to those performed for the BASMI:

- Occiput to wall distance
- Chest expansion

### 8.3.1.9 Enthesitis (MASES)

The MASES comprises 13 items

CCI [REDACTED]

[REDACTED] (Heuft-Dorenbosch et al, 2003) each scored as 0 = yes or 1 = no and then summed for a possible score of 0 to 13, with higher scores indicating worse enthesitis.

If 7 or more items are available, MASES will be imputed by dividing the sum score with the number of assessments and multiplying the result with 13. If less than 7 times are available, MASES will be treated as missing.

Summaries of MASES will be restricted to the subset of subjects with enthesitis present at Baseline. Presence of enthesitis at Baseline will be defined as a Baseline MASES score >0.

#### **8.3.1.10 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 a further unscaled single item (Q2) 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 SF-36 domains (subscores) are scored so that a higher score indicates a better health state.

QualityMetric Health Outcomes™ Scoring Software will be used to calculate norm-based scores for the following domains:

- Physical Functioning (PF)
- Role Physical (RP)
- Bodily pain (BP)
- General health (GH)
- Vitality (VT)
- Social Functioning (SF)
- Role Emotional (RE)
- Mental health (MH)
- Mental Component Summary (MCS)
- Physical Component Summary (PCS)

#### **8.3.1.11 ASAS-NSAID score**

The ASAS-NSAID score is a tool that has been developed to measure the magnitude of NSAID intake during clinical studies (Dougados et al, 2011). Using the concomitant medication NSAID data, the start and stop dates of each medication will be compared to study intervals to determine the dosage and frequency that each NSAID was being taken during the study. Intervals will be weekly from Week 0 to Week 2, bi-weekly up to Week 4 and then every 4 weeks until Week 52.

Subjects not taking NSAIDs in an interval will have an ASAS-NSAID score of 0.

The general formula for calculating the ASAS-NSAID score in an interval is as follows:

- $(\text{equivalent NSAID score}) \times (\text{days of intake during period of interest}) \times (\text{days per week}) / (\text{period of interest in days})$

Each of the components of the above calculation are described below:

- Equivalent NSAID score: This is reported in terms of NSAID equivalent dose in mg/day on a 0–100 scale where the 150mg equivalent diclofenac is set to 100. An NSAID equivalence

table was developed based on a survey of ASAS members. The table including the consensus equivalence score for various NSAIDs can be found in Appendix 13.3.

- Days of intake during period of interest: Equivalent to the total number of days covered during the period that is being measured.
- Days per week: Proportion of days per week when the NSAID is taken. This is collected in the categories described in the days with intake category listed above. Each category corresponds to a score as follows (score in parentheses):
  - Every day or QD (7/7)
  - $\geq 5$  days/week (6/7)
  - 3-5 days/week (4/7)
  - 1-3 days/week (2/7)
  - $< 1$  day/week (0.5/7)
  - No NSAID intake (0)
- Period of interest in days: Refers to the number of days covered for a given NSAID. If only 1 NSAID was taken during the entire period of interest, this will be the same as days of intake during period of interest.

Dougados, et al provide the following example. If during a period of interest (between two visits) of 6 months, the subject has taken piroxicam 20mg during 4 months and if during this 4-month period he has taken piroxicam 3–5 days per week the calculation is as follows:

- $100$  (20mg piroxicam score)  $\times 120$  (4 months)  $\times 4/7$  (3–5 days/week)/ $180$  (6 months) = 38.1

If the subject has used 10mg piroxicam during the remaining 2 months on 2 days a week, the NSAID score for this period is:

- $50$  (10 mg piroxicam score)  $\times 60$  (2 months)  $\times 2/7$  (1–3 days/week)/ $180$  (6 months) = 4.8

In this example the total score for the 6 month period is 42.9 (38.1 plus 4.8).

For the observed case analysis, only intervals that end on or prior to end of the double-blind treatment period will be included. For the LOCF analysis, all data will be included for all intervals up to Week 52 (1 year) irrespective of whether the subject is on double-blind or open-label study medication.

### 8.3.1.12 Work Productivity Survey (WPS)

The WPS is a 9 question instrument used to assess the impact of arthritis on productivity within and outside the home during the preceding 4 weeks. One of the WPS questions concerns the employment status, 3 relate to work productivity outside the home, and 5 ask about household work and daily activities. Subjects employed outside the home will be asked questions about the number of work days missed due to arthritis, the number of days with productivity at work reduced by half or more due to arthritis, and the interference of arthritis on work productivity on a 0 to 10 scale (0=no interference; 10=complete interference). In addition, all subjects regardless of their employment status will be asked questions about the number of household work days missed due to arthritis, the number of days with productivity in household work reduced by half

or more, the number of family, social, or leisure activities days missed due to arthritis, the number of days with outside help hired due to arthritis, and the interference of arthritis on household work productivity on a 0 to 10 scale (0=no interference; 10=complete interference).

In order to make data consistent and amenable to statistical analysis, the following counting rules will be applied to handle out of range and ambiguous answers of the WPS.

These counting rules will be applied prior to conducting any type of statistical analysis of the data.

### WPS counting rules

Due to the inter-relation between certain questions of the WPS, the priority order for implementing these specific counting rules is as in the listed order below.

#### WPS Question 1 (Q1) (employment status)

- If (Q1=missing) and (Q2>0 or Q3>0 or Q4>0), then Q1=YES.
- If (Q1=missing) and {(Q1.a (1) is not missing) or (Q1.a (2) is not missing)}, then Q1=YES.
- If (Q1=missing) and (Q1.b is not missing), then Q1=NO.

For all rules below, the original value will be kept and displayed in the listings but the corrected value will be considered for the analyses (tables).

- If employment status="No" and Job function is not " " then Job function=" ".
- If employment status="No" and occupation is not " " then occupation=" ".
- If employment status="Yes" and subject's status is not " " then status=" ".

#### WPS Q2 (missed work days due to arthritis)

- If (Q1=NO), then Q2="." (missing)
- If Q2 =0 or missing (.) and if Q1 (Employment Status) is Missing (.) then Q2=. (replace 0 by ".").

#### WPS Q3 (days with reduced productivity at work by half or more due to arthritis)

- If (Q1=NO), then Q3="." (missing).
- If Q3 =0 or missing (.) and if Q1 (Employment Status) is Missing (.) then Q3=. (replace 0 by ".").

#### WPS Q4 (rate of arthritis interference on work productivity (work outside of home))

- If (Q1=NO), then Q4="." (missing).
- If Q4 is out of range, then Q4="." (missing).
- If Q4=0 or missing (.) and if Q1 (Employment Status) is Missing (.) then Q4=. (replace 0 by ".").

#### WPS Q9 (rate of arthritis interference on household work productivity)

- If Q9 is out of range, then Q9="." (missing).

## WPS Q2 to Q9

For the rule below, the original value will be kept and displayed in the listings but the corrected value will be considered for the analyses (tables).

- If value=X.5 then value=X+1 (for ex: 1.5->2).

With regards to missing observations at Baseline, the following rules apply:

- Q1 (categorical): missing Baseline is not imputed
- Q2-Q9 (discrete): missing Baseline is imputed by the mean of the available subjects, rounded to the closest integer.

For WPS, the following correction rules have to be applied for Q1, Q2, Q3 and Q4 prior to any analysis of the data to make the data consistent and amenable to statistical analysis:

If Q1='No' and Job function is not ' ' then Job function=' '.

If Q1='No' and occupation is not ' ' then occupation=' '.

If Q1='Yes' and subject's status is not ' ' then status=' '.

If Q1='No' and Q2 is not missing, then Q2='.'

If Q1='No' and Q3 is not missing, then Q3='.'

If Q1='No' and Q4 is not missing, then Q4='.'

### 8.3.1.13 Medical Outcomes Study (MOS) sleep scale

The MOS Sleep Scale is a validated generic self-administered scale measuring specific aspects of sleep. The frequency with which each problem has been experienced during the previous 4 weeks is rated on a 5-point scale ranging from "none of the time" to "all of the time", except sleep quantity, which is reported in hours. All scores are transformed linearly to range from 0 to 100, again with the exception of the sleep quantity subscale, which is scored in hours. Higher scores indicate more of the attribute implied by the scale name (eg, more sleep disturbance, more adequate sleep, or greater sleep quantity).

QualityMetric Health Outcomes™ Scoring Software will be used to calculate the domains of Sleep disturbance, Snoring, Sleep short of breath or headache, Sleep adequacy, Sleep somnolence, Sleep problems Index I, Sleep problems Index II.

### 8.3.1.14 Health Status (EQ-5D)

The EQ-5D consists of a 5-item health status measure and a VAS. Each of the 5 health states is divided into 3 levels (no problem, some or moderate problems, and extreme problems) and is scored as 1, 2, and 3, respectively. The EQ-5D VAS records the respondent's self-rated health status on a vertical 20cm scale, 0 to 100 graduated (0=worst imaginable health status, 100=best imaginable health status).

The VAS will be evaluated by changes from Baseline and actual scores. The 5 dimensions will be analyzed categorically.

As electronic data capture is being used, ambiguous and out of range answers are not expected. However, if these rules are needed, see Appendix 13.2.

### 8.3.1.15 Resources Utilization

The following resource utilization data were collected through UCB standardized modules:

- In-patient hospitalization and emergency room visits
- Health care provider consultations not foreseen by the protocol
- Concurrent medical procedures

Summary statistics and frequency distribution (number of subjects by number of medical resources used) will be presented for resource use with onset during the double-blind period for concurrent medical procedures, health care provider consultations, hospital visits and emergency room (ER) visits. ER visits will be extracted from the hospitalizations/emergency room visit eCRF page (if initial entry point emergency room is ticked).

The categories to be displayed for the frequency distributions of the resource use variables will be defined during the DEM meeting based on blinded data review.

A medical resource is allocated only once in its period of onset, as determined by the (start) date of the event. A resource will be attributed in the same way as an AE is considered a treatment-emergent adverse event (TEAE). The same rules for (partially) missing start and end dates as for AEs will be applied for the resource use.

For the same subject only one hospitalization will be considered if “start date of the second hospitalization - end date of the first hospitalization  $\leq 1$ ”. For missing end dates of hospitalizations (missing discharge dates) the following rules will be applied:

If there is a subsequent hospitalization:

- If the initial entry point and relationship are the same between the two hospitalizations, then the first hospitalization will be grouped with the next one and counted as one hospitalization. In this case the length of the hospitalization will be computed as follows:
  - If the discharge date of the 2<sup>nd</sup> hospitalization is non-missing, then  
length of stay = 2<sup>nd</sup> hospitalization discharge date – 1<sup>st</sup> hospitalization admission date + 1;
  - Else, length of stay = last non missing visit date – 1<sup>st</sup> hospitalization admission date + 1
- Otherwise, if either the entry point or the relationship of the two hospitalizations is different, then both hospitalizations will be counted as distinct. In this case the length of the 1<sup>st</sup> hospitalization will be computed as follows:
  - Length of stay = (2<sup>nd</sup> hospitalization admission date - 1) – 1<sup>st</sup> hospitalization admission date + 1 (ie, 2<sup>nd</sup> hospitalization admission date – 1<sup>st</sup> hospitalization admission date)

In case there is no subsequent hospitalization, then the length of the hospitalization is calculated as length of stay = last non missing visit date – hospitalization admission date + 1

In case of complete consultation date, count only once for a same subject, same consultation date, same location and same provider.

If the procedure name, start date and relationship are the same, then only one procedure is counted, otherwise if at least one variable among procedure name, start date or relationship is different, distinct/several procedures are counted.

If concomitant medical procedures, health care provider consultations not foreseen by the protocol, and hospitalizations/emergency room visits are not available during the study, the respective number variable will be set to 0.

### **8.3.1.16 Extra-articular assessments**

The evolution of associated nonmusculoskeletal features including inflammatory bowel disease, psoriasis, uveitis (including their severity), and flare rate history will be assessed.

## **8.4 Analysis of other efficacy variables during SFE**

In order to assess long-term efficacy of CZP for subjects with nr-axSPA, the following summaries will be presented on the SFE-FAS:

- The number and percentage of subjects achieving response for all categorical variables collected at Week 156/SFE-WD (ASDAS-MI reduction of  $\geq 2.0$  or lowest ASDAS score possible, ASDAS-MI reduction of  $\geq 2.0$ , ASDAS-CII, ASAS40, ASAS20, ASAS partial remission, BASDAI 50 and SPARCC score  $< 2$ ). In addition, a 2-sided exact (Clopper-Pearson) 95% confidence interval for the percentage of subjects achieving response will be presented. As described in Section 4.2.3, the summary will include a composite endpoint analysis, observed case analysis at Week 156 and observed case analysis at Week 156/SFE-WD.
- Subgroup analysis of ASDAS-MI responders (reduction of  $\geq 2.0$  or lowest score possible) and ASAS40 responders by Age group, Gender, Symptom Duration, HLA-B27 Genotype and MRI/CRP classification (as defined in section 4.8) using the composite endpoint analysis described in Section 4.2.3 including 2-sided exact (Clopper-Pearson) 95% confidence intervals for the percentage of subjects achieving response.
- The number and percentage of subjects with ASDAS-ID, ASDAS-MD\*, ASDAS-HD, ASDAS-vHD at Week 156 (completers only) and at Week 156/SFE-WD  
\* ASDAS-MD is equivalent to ASDAS-LD. Whereas protocol amendment 5 uses ASDAS-LD throughout the SAP and Shells will continue to use ASDAS-MD in order to maintain consistency with the double-blind TFLs previously produced for this study.
- Summary statistics along with the changes from baseline for ASDAS, BASDAI (total score, individual Questions 1, 2, 3, 4 and morning stiffness [the average of Q5 and Q6 of the BASDAI]), BASFI, SI Joint SPARCC, PGADA, nocturnal spinal pain, total spinal pain, CRP and ASQoL, collected at Week 156/SFE-WD. As described in Section 4.2.3, observed data for completers only (at Week 156) in addition to observed data for all subjects (completers and subjects with early withdrawal visits) will be presented.

## **9 PHARMACOKINETICS, PHARMACODYNAMICS, AND PHARMACOGENOMICS**

### **9.1 Pharmacokinetics**

Certolizumab pegol plasma concentration data will be summarized by treatment group overall, and by Anti-CZP antibody (ADAb) status of individual subjects (as defined below), for each scheduled visit at which samples were taken, for the SS. Samples collected out of window may

be excluded from the analysis. All CZP concentrations will be reviewed after unblinding by study statistician and clinical pharmacologist to ensure all values are plausible, and values deemed anomalous will be excluded from summary statistics. For concentration data collected in the open-label period, subjects randomized to placebo will only have their OL-CZP data on or after Week 12 open-label visit presented, to ensure the subject is at steady state. For subjects randomized to CZP, their double-blind and OL-CZP data will be presented at all available visits. Geometric mean, geometric coefficient of variation (CV), 95% CIs, arithmetic mean, arithmetic SD, median, minimum, and maximum, will be presented. Geometric mean CZP plasma concentration time curves will be plotted by treatment group, and by antibody status for subjects receiving CZP.

The value of blood sample measurements that are deemed to be below the level of quantification (BLQ), will be set to half the lower level of quantification (LLOQ) for analysis purposes. The summary statistics will only be displayed if at least two-thirds of the values are above the LLOQ.

Individual subject plasma concentrations for CZP, anti-CZP antibody (ADAb) titres, ASDAS scores and dose levels versus time will be produced on the same graph. Two Y-axes on the left will represent the CZP concentrations and anti-CZP levels. The Y-axis on the right will show the ASDAS score. The x-axis will represent time (weeks) with lines showing the treatment and dose.

## 9.2 Pharmacodynamics

Not applicable.

## 9.3 Pharmacogenomics, biomarkers and cytokines

Pharmacogenomics, biomarkers and cytokine data analysis will be described in a separate analysis plan and report.

# 10 IMMUNOGENICITY

## 10.1 Available data

Anti-CZP antibodies (ADAb) will be measured and summarized at Baseline, Weeks 1, 2, 4, 12, 24, 36, 52/WD, and the Follow-up visit (if applicable). In addition, the ADAb will be assessed in subjects who withdraw from study drug and transition to OL-CZP.

A cut point will be determined by the bioanalytical laboratory during assay validation. This cutpoint will be used to determine the status of ADAb in the test sample as above the cut point (ACP) or below the cut point (BCP). For any ADAb test samples with results that are ACP, a further confirmatory assay will be performed, the results of which will be determined as either 'confirmed positive' (CP) or 'not confirmed positive' (NCP).

The following definitions will be applied regarding classification of test samples:

- An ADAb status will be confirmed as positive for any sample with an ADAb level that is ACP and CP
- An ADAb status of negative will be concluded for any sample with an ADAb level that is either BCP or ACP and NCP

Confirmed positive samples will be titrated. The dilution factor will be reported. The titer represents the last dilution factor of the sample's titration series still scoring positive in the screening ADA<sub>b</sub> assay.

In a subset of samples, the ADA<sub>b</sub> response will be characterized for their neutralizing potential. The subset will be selected according to which samples have drug levels to allow testing, and will cover baseline, the visit with the highest titre of ADA<sub>b</sub> on treatment, and follow up. Selection methods will be specified in the bioanalytical plan. A cut point will be determined by the bioanalytical laboratory. This cutpoint will be used to determine the status of neutralizing ADA<sub>b</sub> in the test sample as below the cutpoint (nADA positive) or above the cutpoint (nADA negative).

## 10.2 Subject classification

Subjects will receive an overall classification, inclusive of Baseline and Post-Baseline results, and be classified as follows based on the ADA<sub>b</sub> assay results:

- ADA<sub>b</sub> negative: no confirmed positive ADA<sub>b</sub> samples at any of the sampling time points
- ADA<sub>b</sub> positive: confirmed positive ADA<sub>b</sub> samples at one or more sampling time points
- Missing: relevant samples are missing

Subjects baseline classification will be based on the ADA<sub>b</sub> assay test results at baseline

- Pre-ADA<sub>b</sub> negative: negative ADA<sub>b</sub> baseline sample
- Pre-ADA<sub>b</sub> positive: confirmed positive ADA<sub>b</sub> baseline sample
- Missing: no baseline ADA<sub>b</sub> sample

Baseline is defined as the sample immediately prior to or on the same day as first treatment with CZP. For subjects randomized to CZP, baseline will be Week 0 of the randomized treatment period and for subjects who were randomized to placebo but then received OL-CZP, baseline will be Week 0 of escape OL-CZP.

Subjects will receive a treatment-emergent classification based on the combination of the post-baseline ADA<sub>b</sub> assay results as well as the baseline ADA<sub>b</sub> sample result

- Treatment emergent ADA<sub>b</sub> negative: (i) subjects with no confirmed positive ADA<sub>b</sub> samples at any of the sampling time points, (ii) pre-ADA<sub>b</sub> positive subjects with all post baseline samples either ADA<sub>b</sub> negative or confirmed positive ADA<sub>b</sub> but with a titre below a pre-defined fold increase from the Baseline value (the fold increase from Baseline required to meet these criteria will be defined with the development of the assay and will be included in the TFLs)
- Treatment emergent ADA<sub>b</sub> positive: (i) pre-ADA<sub>b</sub> negative subjects with one or more confirmed positive samples post baseline, (ii) pre-ADA<sub>b</sub> positive subjects with one or more confirmed positive ADA<sub>b</sub> sample post baseline with a titre above a pre-defined fold increase from the Baseline value (the fold increase from Baseline required to meet these

criteria will be the same as that defined for treatment emergent ADA<sub>b</sub> negative and will be included in the TFLs)

- Missing: relevant samples are missing

In addition to the ADA<sub>b</sub> classifications, subjects will also receive an overall nADA classification, inclusive of baseline and post-baseline results, on the nADA assay results

- nADA negative: no nADA positive samples at baseline or post-baseline
- nADA positive: one or more positive samples at baseline or post-baseline
- Missing: relevant nADA samples are missing, e.g. if subject had samples selected for nADA testing based on their ADA<sub>b</sub> levels, but there was insufficient sample left for nADA testing.
- Not Applicable (NA): If a subject was ADA<sub>b</sub> negative (i.e. had no positive ADA<sub>b</sub> samples)

### 10.3 Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by subject.

All analyses will be run on the safety population, unless specified otherwise. ADA<sub>b</sub> and nADA data will be reviewed after unblinding to ensure all values are plausible, and values deemed anomalous will be excluded from summary statistics.

The number and percentage of each subject in the above classifications (ADA<sub>b</sub> and nADA) will be reported. The prevalence of immunogenicity will be reported per time point, defined as (cumulative) proportion of subjects having confirmed positive ADA<sub>b</sub> samples at any point up to and including that time point. Missing samples will not be included in the denominator.

Time to achieving treatment-emergent ADA<sub>b</sub> positivity will be analyzed based on Kaplan-Meier approach, subjects will be considered to have an event at time where treatment emergent ADA<sub>b</sub> positive is first achieved. Subjects classified as treatment-emergent ADA<sub>b</sub> negative will be censored at time of last available ADA<sub>b</sub> result. Discontinued subjects prior to Week 52 will be censored at the last ADA<sub>b</sub> result prior to discontinuation if they have not become ADA<sub>b</sub> Positive. Subjects will be summarized based on completing 52 weeks of treatment, discontinuing prior to Week 52, PBO->OL CZP and CZP->OL CZP treatment groups.

Immunogenicity data will be correlated with PK and efficacy endpoints. No formal inferential statistics (p-values) will be derived. The following types of outputs will be produced:

- Summary of ASDAS-MI responders at Week 12 and 52 as a function of ADA<sub>b</sub> titer. This will be repeated for ASAS40 responders at Week 12 and 52. This will also be presented graphically.
- Spaghetti plots of ADA<sub>b</sub> titre (Y-axis) against time (X-axis) for treatment emergent ADA<sub>b</sub> positive subjects and treatment emergent ADA<sub>b</sub> negative subjects separately for CZP and Placebo (i.e. 4 plots). The CZP plots will have different line patterns for W52 completers and W52 discontinuers.

- Box plots of all (valid) CZP concentrations (Y-axis) versus ADA b titer (X-axis), Y-axis presented in linear and logarithmic scale.

Immunogenicity will also be correlated with possible safety findings.

## 11 SAFETY ANALYSES

The SS will be utilized for safety analyses.

### 11.1 Extent of exposure

There are 4 different concepts for calculating exposure during the double-blind treatment period.

1. This approach will look at the number of doses received, which is defined as dosing days.

Days with 2 injections (i.e. at Week 0, 2 and 4) are counted as 1 dosing day if either injection is given).

2. For the CZP 200mg and placebo treatment arms, duration of exposure to study medication will be calculated as:

Date of last administration of double-blind CZP 200mg or placebo study medication – date of first administration of CZP 200mg or placebo study medication + 14 days (14 days are included in this definition as this is the dosing interval for maintenance of subjects).

If a subject dies during the exposure period (first injection to last injection + 14 days), the exposure period will end with the death date.

For subjects discontinuing from the study before Week 52, the 2-week period after last double-blind injection will be utilized.

3. For the study exposure of a subject, 5 half-lives of CZP will be taken into account. Hence, a subject will be regarded as being exposed to study drug from first double-blind injection to last double-blind injection + 70 days. The days of drug holidays beyond 70 days will be subtracted.

If a subject dies during the exposure period (1<sup>st</sup> injection to last injection + 70 days), the exposure period will end with the death date.

For subjects discontinuing from the study before Week 52, the 70 day period after last injection will be utilized.

4. This approach will be the nearly identical to approach 2. However, exposure will be censored at the first occurrence of the AE to be considered for analysis (a separate calculation has to be performed for each preferred term). The different exposure duration for the respective AEs will only be displayed in the AE tables for exposure-adjusted incidences and will not be summarized in the exposure table.

Exposure to the OL-CZP treatment will also be calculated using the above 4 methods considering only dosing during the time the subject is receiving OL-CZP study medication prior to Start of SFE. Exposure during the SFE period will be calculated using methods 2, 3 and 4 assuming first dose in the SFE period starts when the subject takes their loading dose/first dose of OL-CZP in the SFE period and ends when the subject has their last administration of study medication (or if missing, the date of withdrawal from the SFE). Total CZP incl DB and OL

exposure will be calculated using all 4 methods, by summing the exposure to CZP during the double-blind and OL- CZP periods. Total CZP incl DB, OL, and SFE OL will be calculated using methods 2, 3 and 4, by summing the exposure to CZP during the double-blind, OL-CZP and SFE periods. Note that for Total CZP incl DB, OL, and SFE OL method 1 will not be calculated as the SFE period does not collect data on dose administration. Exposure to OL-OT will not be calculated.

For the first 3 approaches, tables will summarize exposure for the placebo, CZP 200mg Q2W, OL-CZP treatment groups (split by the treatment taken during the double-blind period), SFE OL-CZP, Total CZP incl DB and OL and Total CZP incl DB, OL, and SFE OL (for methods 2 and 3 only).

The following doses will be presented in a listing:

- very first dose in the study, the very last dose in the study and the difference between them
- last dose in the DB period (i.e. last DB dose or last OL dose), the last SFE dose and the difference between them.

## 11.2 Adverse events

AE start dates will be examined to determine if they are treatment-emergent and attributable to the double-blind period, open-label period, pre-SFE period or SFE period. AEs occurring either before double-blind study drug administration (pre-treatment) or outside of the windows described below will be defined as non-treatment emergent.

**Table 11–1: Assignment of Adverse Events to Study Periods**

Type of Subject	Assignment Rule Based on AE Start Date	Assignment
Subjects who do not enter the open-label period and do not enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days.	Double-blind TEAE
Subjects who do not enter the open-label period but do enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and prior to the SFE period start	Double-blind TEAE

**Table 11–1: Assignment of Adverse Events to Study Periods**

Type of Subject	Assignment Rule Based on AE Start Date	Assignment
Subjects who enter the open-label period	<p>1. Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and before the start date and time of OL-CZP or OL-OT.</p> <p><b>Note:</b> If a subject continues in the study with OL-OT visits but without taking any relevant other medication, there will be no start date of the other treatment period and all AEs with a start date within 70 days after last administration will be assigned to the DB period.</p>	Double-blind TEAE
	<p>2. Start date is on or after the start date of OL-CZP and on or before the last study drug administration in the OL-CZP period + 70 days and</p> <ul style="list-style-type: none"> <li>• before the start date of OL-OT (for subjects moving into OL-OT), or</li> <li>• before the SFE period start (for subjects moving into SFE period)</li> </ul>	Open-label CZP TEAE
	<p>3. Start date is on or after the start date of OL-OT.</p> <p><b>Note:</b> Other treatment adverse events will not be summarized, but only presented in listings</p>	Open-label other treatment TEAE
Subjects who enter the pre-SFE period	Start date is after the last double-blind or OL-CZP dose + 70 days and is before the SFE period start.	Pre-SFE AE
Subjects who enter the open-label SFE period	Start date is on or after the SFE period start and on or before the SFE period end + 70 days.	SFE TEAE

If AE intensity or relationship is missing, then the given event will be imputed as severe or related, respectively.

The frequency of all AEs including TEAEs will be presented for each treatment group separately by System Organ Class, high level term, and preferred term. The data will be displayed as number of subjects experiencing the AEs, percentage of subjects, and number of AEs. Certain summaries will also include exposure-adjusted incidence rates and event rates. TEAEs will be summarized by treatment and period as described in [Section 3.6](#).

The following AE summaries will be presented:

- Overview of TEAE
- Overview of Pre-SFE AEs

- TEAE
- Serious TEAE
- Non-serious TEAE
- TEAE with fatal outcome
- TEAE leading to permanent withdrawal of study medication
- Severe TEAE (included in TEAE table by intensity)
- Drug-related TEAE (included in TEAE table by relationship)
- Serious TEAE by relationship
- Non-serious TEAE by relationship
- Fatal TEAE by relationship
- TEAE of special interest (see below)
- TEAE sorted by incidence of PTs in the CZP 200mg group (SOC and higher level term [HLT] will not be utilized)
- Non-serious TEAE with an incidence of more than 5% in 1 of the 2 randomized groups (HLT will not be utilized)
- Non-serious TEAE with an incidence of more than 5% in 1 of the 2 randomized groups by relationship (HLT will not be utilized)
- TEAE subject numbers
- TEAEs with a CZP incidence  $\geq 1\%$  and CZP incidence  $>$  PBO incidence

The AEs of interest and the approach for summarizing them is briefly described below. Further information is provided in the guidance document (AEs of Interest – Cimzia Program 2018-01-05).

1. Serious infections, including opportunistic infections. Serious infections will be summarized using the Serious TEAE table. No separate table is planned. In addition, opportunistic infections (including tuberculosis) will be presented in a table using UCB-defined search criteria.
2. Malignancies, including lymphoma. These will be presented in 2 tables using the Standardized MedDRA Query (SMQ) criteria “Malignant or unspecified tumours” and “Malignant tumours” respectively.
3. Serious cardiovascular events (also called major adverse cardiac events or MACE). These will be presented in a table including all serious TEAEs with the SMQs of “Haemorrhagic central nervous system vascular conditions”, “Conditions associated with central nervous system haemorrhages and cerebrovascular accidents”, and “Ischaemic central nervous system vascular conditions” with the exception of events coding to a PT of “Transient ischaemic attack”, including all serious TEAEs with the HLT of “Ischaemic coronary artery disorders” except events coding to PTs of “Chest Pain” or “Chest discomfort” and including all serious

TEAEs with HLTs of “Heart Failures NEC”, “Left Ventricular Failures”, or “Right Ventricular Failures” and which also code to the SOC of “Cardiac Disorders”.

4. Congestive heart failure. These will be manually identified by the study physician from the TEAE table. No separate table is planned.
5. Demyelinating-like disorders will be presented in table which is based on the SMQ = “Demyelination”. The SMQ search should include all TEAEs which code to a PT included in Scope=Narrow group within the SMQ. TEAEs which code to a PT included in the Scope=Broad group within the SMQ should be excluded from the search.
6. Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia. These will be presented in a table using the criteria SMQ = “Haematopoietic cytopenias” in the subset of Serious TEAEs.
7. Serious bleeding events. These will be presented in a table using the criteria SMQ = “Haemorrhage terms (excl laboratory terms)” in the subset of Serious TEAEs.
8. Lupus and lupus-like illness. These will be manually identified by the study physician from the TEAE table. No separate table is planned.
9. Serious skin reactions (eg, Stevens-Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme). These will be manually identified by the study physician from Serious TEAE table. No separate table is planned.

The tables for 1, 2, 3, 5, and 6 above, will include the exposure-adjusted incidence rate (EAIR) with associated 95% CI, and the exposure-adjusted event rate (EAER) as described below.

Although not officially AEs of interest, the following events will also be summarized as described in the AEs of Interest – Cimzia Program 2018-01-05 guidance document.

1. Hepatic events. These will consist of a subset of all TEAEs, identified using the following 5 SMQs: Cholestasis and jaundice of hepatic origin; Hepatic failure, fibrosis, and cirrhosis and other liver damage-related conditions; Hepatitis, noninfectious; Liver-related investigations, signs and symptoms; and Liver-related coagulation and bleeding disturbances.
2. Hypersensitivity reactions and anaphylactic reactions. Hypersensitivity reactions will be summarized as any TEAEs occurring on the same day or the day after injection was received, which code to the following preferred terms: Administration site hypersensitivity, Documented hypersensitivity to administered product, Drug hypersensitivity, Hypersensitivity, Hypersensitivity vasculitis, Infusion site hypersensitivity, Injection site hypersensitivity, Medical device site, hypersensitivity, Type II hypersensitivity or Type IV hypersensitivity reaction. Anaphylactic reactions will be defined using an algorithmic approach as described in the “AEs of Interest – Cimzia Program 2018-01-05” guidance document.

Because subjects will be able to adjust background medications during the study, some additional AE tables may be prepared in which events occurring while on certain background medications will be summarized. The list of these supportive tables is expected to be as follows but other subsets of TEAEs may be explored:

- TEAEs reported after starting a new biologic, but within 10 weeks of stopping CZP

- TEAEs reported one month after changing NSAID type
- TEAEs reported three months after changing or adding a DMARD
- TEAEs reported within one month of an addition of oral corticosteroids or 50% increase in dose

The purpose of these tables will be to assess whether any observed treatment differences in the AE profile may be due, in part, to these prespecified background medications.

A TEAE table will also be presented for the ‘anti-CZP antibody status’ subgroup defined in [Section 10](#). Only subjects exposed to CZP and only TEAEs occurring after first dose of CZP will be considered. A further column will be presented summarizing TEAEs occurring after the onset of the positive antibody status. The EAIR will be presented in addition to the associated 95% CI, and the EAER.

The TEAE incidence rates will also be adjusted for exposure (as calculated using methods 3 and 4 in [Section 11.1](#)) and reported by 100 subject years of exposure. Two approaches will be applied to adjust for exposure. One will only use the first occurrence of an AE with corresponding exposure (EAIR) and the second approach will use all AEs and the entire exposure (EAER).

For the EAIR, the first occurrence of an TEAEs for a certain treatment group will be utilized (either placebo or CZP 200mg Q2W for the double-blind period or OL-CZP) and divided by the sum of exposure of all subjects in the respective treatment group, where subjects experiencing the respective AE will be censored at the time of occurrence of the AE. The EAIR will be multiplied by a factor of 100 to give a rate per 100 subject-years.

For the EAER, all TEAEs for a certain treatment group will be utilized (either placebo or CZP 200mg Q2W for the double-blind period or OL-CZP) and divided by the sum of exposure of all subjects in the respective treatment group. The EAER will be multiplied by a factor of 100 to give a rate per 100 subject-years.

The summary of TEAEs, serious TEAEs, and TEAEs leading to permanent withdrawal of study medication will include the EAIR (with corresponding 95% exact CI) and EAER. The confidence interval for the EAIR will be based on Poisson rate CIs. The bounds of the exact 95% CIs are estimated by

$$\text{lower bound} = \frac{1}{2 \cdot t} \cdot \chi^2_{2 \cdot (\text{number of patients with events}), 0.025} \quad \text{and} \quad \text{upper bound} = \frac{1}{2 \cdot t} \cdot \chi^2_{2 \cdot (\text{number of patients with events} + 1), 0.975}$$

$\chi^2_{n,\gamma}$  denotes the chi<sup>2</sup>-distribution with n degrees of freedom and quantile  $\gamma$  and t is the exposure censored at the time of occurrence of the AE.

### Date imputation for incomplete or missing start and/or stop dates

Where an AE start date is (partially) missing, the AE will be considered TE if possible.

Imputation rules provided in [Section 4.2.4](#) 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.

In order to identify differences between CZP and placebo with respect to TEAEs, the relative risk (CZP/placebo) of the most common TEAEs (incidence >1% for the PT in the CZP treatment group) will be presented graphically by descending relative risk including the 95% CI.

Listings for TEAEs, serious TEAEs, TEAEs leading to withdrawal of study medication, and TEAEs with fatal outcome will be provided. A glossary for the reported terms will also be generated utilizing SOC, HLT, and PT.

### 11.3 Clinical laboratory evaluations

The changes from Baseline in laboratory evaluations will be analyzed over time for the SS. In addition, last value (end of treatment), minimum value during treatment, and maximum value during treatment will be analyzed by period (double-blind and OL-CZP). End of treatment will be defined as last visit not including the Follow-up visit. During treatment for the minimum and maximum calculation will also exclude the Follow-up visit.

Shift tables concerning the normal range at end of treatment, minimum and maximum shift at any time will also be produced for each hematology and biochemistry laboratory parameter. The shifts will be categorized using L, N, H, missing, and total.

The number and percent of subjects with markedly abnormal ( $\geq$ grade 3 by RCTC) hematology or biochemistry values will be summarized at Any Visit, On Treatment (excluding the Follow-up visit) and by visit separately by period (double-blind or OL-CZP). Subject numbers for subjects with any markedly abnormal hematology or biochemistry value will be tabulated. Values fulfilling the criteria below will be classified as marked abnormal high (MH) or marked abnormal low (ML). If no lower (upper) limit is given, the classification ML (MH) is not applicable.

- Hemoglobin < LLN and decrease from Baseline >2g/dL
- Hemoglobin <8g/dL
- White blood cells <2000/ $\mu$ L
- Lymphocyte count <1000/ $\mu$ L
- Neutrophil count <1000/ $\mu$ L
- Platelet count <50000/ $\mu$ L
- ALT >3x upper limit normal (ULN)
- AST >3x ULN
- ALP >3x ULN
- Bilirubin  $\geq$ 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  $\geq$ 3x ULN

The liver function test elevations will be displayed in an incidence table for the entire double-blind treatment period (including the Follow-up visit) and separately for the OL-CZP treatment period. A shift table will also be presented for double-blind and OL-CZP periods separately showing the change from baseline to maximum post-baseline category ( $\leq$ 1xULN, >1-<2xULN, 2-<3xULN and  $\geq$ 3xULN).

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.

## 11.4 Vital signs, physical findings, and other observations related to safety

### 11.4.1 Vital signs

The changes from Baseline in vital signs will be analyzed over time for the SS. In addition, the last value (end of treatment), minimum value during treatment, and maximum value during treatment will be analyzed by period (double-blind and OL-CZP). End of treatment will be defined as last visit not including the Follow-up visit. During treatment for the minimum and maximum calculation will also exclude the Follow-up visit.

Early Withdrawal Visits will be assigned to what would have been the next scheduled visit.

The number and percent of subjects with abnormal vital sign values will be summarized by visit and for any visit. Abnormal vital sign results are defined as follows.

<b>Vital Sign Parameter</b>	<b>Abnormally Low</b>	<b>Abnormally High</b>
Systolic Blood Pressure (mmHg)	$\leq$ 90 and decrease of $\geq$ 20	$\geq$ 180 and increase of $\geq$ 20
Diastolic Blood Pressure (mmHg)	$\leq$ 50 and decrease of $\geq$ 15	$\geq$ 105 and increase of $\geq$ 15
Pulse Rate (bpm)	$\leq$ 50 and decrease of $\geq$ 15	$\geq$ 120 and increase of $\geq$ 15

## **11.4.2 Electrocardiograms**

Not applicable.

## **11.4.3 Other safety variables**

### **11.4.3.1 Weight**

The changes from Baseline in weight will be analyzed over time in the SS.

### **11.4.3.2 Pregnancy testing**

Pregnancy testing will consist of serum testing at Screening and Follow-up and urine testing (dipstick) at Baseline, Week 0 of the alternative study assessment for subjects administering either OL-CZP or OL-OT) and at Week 52/WD. No tables will be generated and data will be listed only.

### **11.4.3.3 Physical assessments**

Physical examination will be performed at Screening, Baseline, Weeks 2, 4, 8, 12, 16, 24, 36, 52/WD and at the Follow-up visit (8 weeks after the Week 52/WD visit for subjects not participating in the SFE period). Physical examination findings will be recorded in the eCRF only at Screening. Physical examination data will be listed only.

### **11.4.3.4 Tuberculosis assessments**

The TB assessments at Screening will only be performed to verify the eligibility criteria (presence of active TB or new latent TB infection). These data will be presented as Baseline characteristics (see [Section 6.2](#)).

The TB assessments during the study will be utilized to check if a subject has to be withdrawn from the study. QuantiFERON-TB Gold Tuberculosis Test Results will be presented by Visit including End of Treatment.

## 12 REFERENCES

Anderson JJ, Baron G, van der Heijde D, Felson DT, Dougados M. Ankylosing spondylitis assessment group preliminary definition of short term improvement in ankylosing spondylitis. *Arthritis Rheum*. 2001 Aug;44(8):1876-86.

Brandt J, Listing J, Sieper J, Rudwaleit M, van der Heijde D, Braun J. Development and preselection of criteria for short term improvement after anti-TNF $\alpha$  treatment in ankylosing spondylitis. *Ann Rheum Dis*. 2004; 63:1438-44.

Dougados M, Paternotte S, Braun J, Burgos-Vargas R, Maksymowych W, Sieper J, van der Heijde D. ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. *Ann Rheum Dis* 2011;70:249–251.

Doward LC, Spoorenberg A, Cook SA, Whalley D, Helliwell PS, Kay LJ, et al. Development of the ASQL: a quality of life instrument specific to ankylosing spondylitis. *Ann Rheum Dis*. 2003; 62:20-6. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, Landewé R, van der Tempel H, Mielants H, et al. Assessment of enthesitis in ankylosing spondylitis. *Ann Rheum Dis*. 2003;62:127-132.

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

Machado P, Navarro-Compán V, Landewé R, van Gaalen FA, Roux C, van der Heijde D. How to calculate the ASDAS if the conventional CRP is below the limit of detection or if using high sensitivity CRP? – An analysis in the DESIR cohort. *Arthritis Rheum*. 2014 Accepted Article; doi: 10.1002/art.38921.

Machado PM, Landewé R, van der Heijde D. Assessment of SpondyloArthritis international Society (ASAS). Letter. Ankylosing Spondylitis Disease Activity Score (ASDAS): 2018 update of the nomenclature for disease activity states. *Ann Rheum Dis*. 2018;77:1539-40.

Maksymowych WP, Mallon C, Richardson R, et al. Development and Validation of the Edmonton Ankylosing Spondylitis Metrology Index. *Arthritis Rheum*. 2006;55:575-82,

O’Kelly M, Ratitch B. *Clinical trials with missing data: A guide for practitioners*. (1st Ed). 2014,

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, Lie E, Kvien TK, Sieper J, Van den Bosch F, Listing J, et al. The ASDAS is a highly discriminatory ASAS-endorsed disease activity score in patients with ankylosing spondylitis. *Ann Rheum Dis*. 2008;doi:10.1136/ard.2008.100826.

Ware JE, Kosinski M, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. *User’s manual for the SF-36 v2 Health Survey* (2nd Ed). 2007.

---

**13 APPENDICES**

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

### 13.1 Handling of questionnaire data

The following rules will apply for analysis of (1) out of range and (2) ambiguous answers (ie, invalid or unable to interpret answers) to questionnaires completed by subjects:

In case of out of range answer (ie, an answer that does not correspond to any possible response proposed in the questionnaire, eg, “?”, “I don’t know,” or any value superior or inferior to the ones specified in the response options): the answer will be scored “missing”.

However, in case the subject selected one of the proposed responses but added a comment (for instance “6 +++” or “5 ?”), the response (ie, “6” or “5”) will be retained for scoring but not the comment (ie, “+++” or “?”).

In the same way, if the subject selected one of the proposed responses but added a value superior or inferior to the ones specified in the responses options (for instance, “4/5” or “-1/2” on a 5-point scale ranging from 0 to 4), the response corresponding to the possible responses options (ie, “4” or “2”) will be retained for scoring but not the values superior or inferior to the responses options (ie, “5” or “-1”).

In case of ambiguous answer (ie, multiple responses to a question allowing only a single response, a response marked between two allowed responses):

Multiple responses to a question allowing only a single response:

- If half or more responses are marked (ie, 4 responses marked on a seven point scale, 3 responses marked on a 5-point scale, 2 responses to a Yes/No item...): the answer will be scored “missing”.
- If less than half responses are marked:
  - if the responses are NOT adjacent to each other: the answer will be scored “missing”,
  - if the responses are adjacent to each other (“2/3” or “2/3/4”, for instance), the more severe score will be retained.

If a response is marked between two allowed responses (for instance, the subject marked his/her response between 2 and 3 on a 4-point scale allowing only responses 1, 2, 3 and 4): the nearest more severe score will be retained.

If a response expected to be stable over time (eg, education) is varying over time no corrective action is foreseen and data will be utilized as reported.

If the EQ-5D VAS contains 2 answers, the most severe answer will be retained.

Please see [Section 8.3.1.12](#) for special rules applicable to the WPS.

### 13.2 ASAS-NSAID equivalent score

**Table 13–1: ASAS-NSAID equivalent score**

NSAID	Consensus dose comparable to 150mg of diclofenac (in mg)
Diclofenac	-
Naproxen	1000
Aceclofenac	200
Celecoxib	400
Etodolac	600
Etoricoxib	90
Flurbiprofen	200
Ibuprofen	2400
Indometacin	150
Ketoprofen	200
Meloxicam	15
Nimesulide	200
Phenylbutazone	400
Piroxicam	20
Tenoxicam	20

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

## **14 AMENDMENT(S) TO THE STATISTICAL ANALYSIS PLAN (SAP)**

### **14.1 AMENDMENT 1**

#### **Rationale for the amendment**

The analysis plan was updated to ensure consistency with the substantial protocol amendment 1, which included several changes to clarify and /or add supporting information regarding the procedures and assessments, and to remove inconsistencies and errors. A range of sensitivity analyses were added into the protocol after discussion with the regulatory authorities, to evaluate the impact of missing data on the analysis of the primary efficacy variable.

As the original SAP was finalized to aid regulatory discussion and the subsequent changes were substantial, it was considered not necessary to document in detail all modifications and changes in this section. However, a brief summary of changes are provided below:

- Protocol scheduled assessments updated for consistency with protocol amendment 1
- Safety Set definition updated to match protocol amendment 1
- The use of graphs for individual plasma concentrations for CZP and anti-CZP levels versus time for consistency with protocol amendment 1
- Methods for handling missing data and sensitivity analyses of efficacy endpoints updated for consistency with protocol amendment 1
- Further detail provided on the analysis of data after a subject discontinues from the double-blind period
- Imputation of partial dates rules updated
- Further detail provided on subgroup analyses
- Other changes made in this amendment are to provide clarification or are administrative in nature, including minor editorial changes to abbreviations

### **14.2 AMENDMENT 2**

#### **Rationale for the amendment**

The analysis plan was updated to ensure consistency with protocol amendment 3, which included an additional SFE period, for an additional 2 years, where subjects may receive open-label CZP treatment. All modifications are detailed below.

#### **Change #1**

AS0006 has been replaced with AS0006 (C-AXPAND) on the title page and sections 1, 2 and 2.3.

#### **Change #2**

DEM, SMQ and SFE were included as acronyms. IVRS was corrected to be IXRS throughout.

### **Change #3**

#### **Section 2:**

AS0006 is a randomized, double-blind, parallel-group, placebo-controlled study, consisting of 52 weeks of treatment with certolizumab pegol (CZP) and a follow-up period for 8 weeks after Week 52/Withdrawal [WD] visit which is 10 weeks after the last administration of study medication.

#### **Was changed to:**

The C-axSpAnd study (AS0006) is a randomized, double-blind, parallel-group, placebo-controlled study, consisting of 52 weeks of treatment with certolizumab pegol (CZP), followed by a follow-up period for 8 weeks after the Week 52/Withdrawal [WD] visit which is 10 weeks after the last administration of study medication or a 2 year extension with open-label CZP treatment.

### **Change #4**

#### **Section 2.2.3:**

Anti-CZP antibody (ADAb) concentrations will be measured and summarized at Baseline, Weeks 1, 2, 4, 12, 24, 36, 52/WD, and the FU visit. In addition, the anti-CZP Ab concentrations will be assessed in subjects who withdraw from study drug and transition to open-label CZP. The number and percentage of subjects with anti-CZP antibody concentrations above 2.4 units/mL will be reported as follows:

- Number and percentage of subjects with ADAb > 2.4 units/mL at the time of each visit
- Number and percentage of subjects with ADAb > 2.4 units/mL at any visit during treatment (not including post treatment withdrawal or Follow-up visit)
- Number and percentage of subjects with ADAb > 2.4 units/mL at any visit including post treatment withdrawal or Follow-up visit

#### **Was changed to:**

Anti-CZP antibody (ADAb) concentrations will be measured and summarized at Baseline, Weeks 1, 2, 4, 12, 24, 36, 52/WD, and the Follow-up visit (if applicable). In addition, the anti-CZP Ab concentrations will be assessed in subjects who withdraw from study drug and transition to open-label CZP. The number and percentage of subjects with anti-CZP antibody concentrations above a defined cut point will be reported as follows:

- Number and percentage of subjects with ADAb > defined cut point at the time of each visit
- Number and percentage of subjects with ADAb > defined cut point at any visit during treatment (not including post treatment withdrawal or Follow-up visit)
- Number and percentage of subjects with ADAb > defined cut point at any visit including post treatment withdrawal or Follow-up visit

The cut point to use will be agreed after the sample analysis at the data evaluation meeting (DEM) prior to database lock and unblinding.

## **Change #5**

The following text added as the last sentence in section 2.3 “At the completion of the Week 52 visit assessments, eligible subjects may receive open-label CZP treatment for an additional 2 years in an open-label, Safety Follow-up Extension (SFE) period.”

## **Change #6**

### **Section 2.3.1:**

Period 2 (Double-blind period) – Week 0 to Week 52, placebo-controlled.

Eligible subjects will be allocated to the following double-blind study treatments in a 1:1 ratio:

- CZP administered sc at the loading dose of CZP 400mg Q2W at Weeks 0, 2 and 4 followed by CZP 200mg Q2W (starting at Week 6)
- Placebo

At on-site visit days (at Weeks 0, 2, 4, 6, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 50), the double-blind study treatments (including placebo) will be administered sc at the study site. At the Week 0, 2, 4, 6, and 8 visits, the administration must be done by dedicated, unblinded, and adequately trained site personnel. At the Week 10 and 12 visits, the subject will receive training on self-injection and the study treatment will be self-administered under the supervision of the unblinded study personnel. At Weeks 16, 20, 24, 28, 32, 36, 40, 44, 48, and 50, the study treatment administration should preferably be done by the unblinded study personnel, or alternatively by self-injection of the subject under the supervision of the unblinded study personnel.

Period 3 (Follow-up period):

All subjects, including those withdrawn from the double-blind study treatment, will have a Follow-up visit, 8 weeks after Week 52/WD visit (10 weeks after their last administration of study medication).

Subjects who discontinue the double-blind study treatment and enter open-label treatment with CZP will be assessed 2 and 4 weeks after open-label CZP treatment has been initiated and every 12 weeks thereafter. Alternatively, subjects who discontinue the double-blind study treatment and receive other treatment (including biologics) will be assessed every 12 weeks after initiation of the other treatment.

The schedule of study assessments is provided in [Section 5.2](#) of the protocol.

A schematic diagram of the study design is provided in [Section 5.3](#) of the protocol.

### **Has been changed to:**

Period 2 (Double-blind period) – Week 0 to Week 52, placebo-controlled.

Eligible subjects will be allocated to the following double-blind study treatments in a 1:1 ratio:

- CZP administered sc at the loading dose of CZP 400mg Q2W at Weeks 0, 2 and 4 followed by CZP 200mg Q2W (starting at Week 6)
- Placebo

At on-site visit days (at Weeks 0, 2, 4, 6, 8, 10, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 50), the double-blind study treatments (including placebo) will be administered sc at the study site. At the Week 0, 2, 4, 6, and 8 visits, the administration must be done by dedicated, unblinded, and adequately trained site personnel. At the Week 10 and 12 visits, the subject will receive training on self-injection and the study treatment will be self-administered under the supervision of the unblinded study personnel. At Weeks 16, 20, 24, 28, 32, 36, 40, 44, 48, and 50, the study treatment administration should preferably be done by the unblinded study personnel, or alternatively by self-injection of the subject under the supervision of the unblinded study personnel.

Subjects who discontinue the double-blind study treatment and enter open-label treatment with CZP will be assessed 2 and 4 weeks after open-label CZP treatment has been initiated and every 12 weeks thereafter. Alternatively, subjects who discontinue the double-blind study treatment and receive other treatment (including biologics) will be assessed every 12 weeks after initiation of the other treatment. In either alternative schedule, assessments should be conducted until as close as possible to Week 52 (within  $\pm 4$  weeks), where Week 52 is relative to the original randomization at Week 0. Subjects will then be invited to the Week 52 assessment visit.

Period 3 (Follow-up period) - All subjects not participating in the SFE period after Week 52, including those withdrawn from the study prematurely, will have a single Follow-up visit, 8 weeks after Week 52/WD visit (10 weeks after their last administration of study medication).

SFE Period - Week 52 to Week 156, open-label:

At the completion of the Week 52 visit assessment, subjects on double-blind study treatment may receive open-label CZP treatment for an additional 2 years. Subjects who withdrew from double-blind study treatment and transitioned to open-label CZP treatment are also eligible to participate in the SFE period after completing the Week 52 visit assessment. Subjects on other treatments are not eligible to participate in the SFE period.

Eligible subjects are allowed to roll-over to the SFE-period up to 3 months after completion of the Week 52 assessments.

Starting at Week 64, all subjects are to visit the site approximately every 12 weeks for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156.

The schedule of study assessments is provided in [Section 5.2](#) of the protocol.

A schematic diagram of the study design is provided in [Section 5.3](#) of the protocol.

### **Change #7**

#### **Section 2.3.2:**

For each subject, the study will last a maximum of 66 weeks, as follows:

- Up to 6 weeks in the Screening period
- 52 weeks in the double-blind period
- A Follow-up visit, 8 weeks after Week 52/WD visit, 10 weeks after last dose administration (ie, Week 50, if the subject completed the entire dose administration schedule)

**Has been changed to:**

For each subject, the first 3 periods of the study will last a maximum of 66 weeks, as follows:

- Up to 6 weeks in the Screening period
- 52 weeks in the double-blind period
- A Follow-up visit, 8 weeks after Week 52/WD visit, for subjects not participating in the SFE Period.

For subjects participating in the SFE period, the study will extend up to a maximum of 104 additional weeks.

**Change #8**

**Section 2.3.3:**

The end of the study is defined as the date of the last visit (including Follow-up visit) of the last subject in the study.

**Has been changed to:**

The end of the study will be defined as the date of the last subject's last visit, defined as the Follow-up Visit 8 weeks after Week 52/WD visit for subjects not participating in the SFE Period, or the last visits of the SFE Period.

**Change #9**

**Section 2.3.3:**

The end of the study is defined as the date of the last visit (including Follow-up visit) of the last subject in the study.

**Has been changed to:**

The end of the study is defined as the date of the last visit (including Follow-up visit and SFE period) of the last subject in the study. Period start and ends are defined in [Section 3.2.1](#).

**Change #10**

**Section 2.3.4:**

Approximately 900 subjects are expected to enter the Screening period in order to have 300 subjects randomized into the double-blind period. It is planned to enroll the subjects at approximately 95 sites.

**Has been changed to:**

Approximately 1200 subjects are expected to enter the Screening period in order to have 300 subjects randomized into the double-blind period. It is planned to enroll the subjects at approximately 120 sites.

## **Change #11**

### **Section 2.3.6:**

Subjects who are MRI-/CRP- are not eligible for randomization. The IVRS will be designed to ensure that at least 20% and no more than 40% of the randomized subjects belong to one of the three clinical subgroups above.

### **Has been changed to:**

Subjects who are MRI-/CRP- are not eligible for randomization. The interactive response system (IXRS) will be designed to ensure that at least 20% of the randomized subjects belong to one of the three clinical subgroups above.

## **Change #12**

### **Section 3.2.1.2:**

The double-blind treatment period begins at the Baseline visit (Week 0) at the time of first study medication administration (inclusive) and ends at Week 52/WD visit. The Follow-up visit will take place 10 weeks after the last dose of study medication, which will be 8 weeks after the Week 52 visit (if the Week 52 is the last non Follow-up visit) or less than 10 weeks after the W/D visit (if the W/D visit is the last non Follow-up 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 52 Visit without early withdrawal of the study. This is regardless of whether they attend the Follow-up visit and regardless of whether they are on double-blind treatment, open-label CZP, or other treatment.

### **Has been changed to:**

The double-blind treatment period begins at the Baseline visit (Week 0) at the time of first study medication administration (inclusive) and ends when:

- the subject completes the Week 52 visit on double-blind treatment, for subjects continuing into the Follow-up period
- the subject takes their loading dose/first dose of open-label CZP in the SFE period, for subjects completing the Week 52 visit on double-blind treatment and continuing into the SFE
- the subject takes their first dose of open-label CZP or other treatment (if prior to Week 52 visit)
- the subject discontinues (WD visit) and does not fall into the above categories.

For subjects who prematurely withdraw from the study prior to Week 52 (for example, those who cease to have double-blind and open-label visit data collected), their visit assessments at withdrawal will be assigned to the next scheduled visit following the last visit where assessments are available according to each protocol activity. This could be the next double-blind or open-label visit. 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 52 visit without early withdrawal of the study. This is regardless of whether they attend the Follow-up visit or enter SFE period and regardless of whether they are on double-blind treatment, open-label CZP, or other treatment.

### **Change #13**

#### **Section 3.2.1.3:**

The first day of the open-label period starts when a subject has discontinued the double-blind treatment period and they take their first dose of open-label CZP or other therapy. The open-label period ends when the subject discontinues the study.

#### **Has been changed to:**

The first day of the open-label period starts when a subject has discontinued the double-blind treatment period prior to Week 52 and they take their first dose of open-label CZP or other treatment. The open-label period ends when:

- the subject completes the Week 52 visit on open-label CZP or other treatment, for subjects continuing into the Follow-up period
- the subject takes their loading dose/first dose of open-label CZP in the SFE period, for subjects continuing into the SFE
- the subject discontinues the study (WD visit) and does not fall into the above categories.

### **Change #14**

#### **The following two sections have been added:**

##### **3.2.1.4 Follow-up period**

For patients not entering the SFE period (no SFE informed consent), the Follow-up period will start on the day after the Week 52 visit (for subjects having Week 52 visit on double-blind or open-label therapy) or on the day after the WD visit. The Follow-up period will end on the Follow-up visit date. The Follow-up visit will take place 10 weeks after the last dose of study medication, which will be 8 weeks after the Week 52 visit (if the Week 52 is the last non-Follow-up visit) or less than 10 weeks after the WD visit (if the WD visit is the last non-Follow-up visit).

##### **3.2.1.5 SFE period**

The SFE period starts when the subject takes their loading dose/first dose of open-label CZP in the SFE period. The SFE period ends when the subject discontinues the study.

### **Change #15**

#### **Section 3.2.2:**

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 double-blind 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 double-blind 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 '+’.
- 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.

**Has been changed to:**

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

- If the start (stop) date occurred on or after the first dose, and on or prior to the double-blind drug stop date, relative day is calculated as start (stop) date minus first double-blind dose date + 1. Relative day will be prefixed by a 'd' in the listings to show it's relative to double-blind treatment.
- For patients not entering the open-label CZP period (including SFE period), if the start (stop) date occurred after the last dose of double-blind drug, the relative day to the most recent double-blind dose is calculated as start (stop) date minus most recent double-blind dose date. The relative day in this situation should be preceded by a 'd+'.
- For subjects entering the open-label CZP period (including SFE period), if the start (stop) date occurred on or after the first open-label CZP start date, and on or prior to the open-label CZP stop date (including SFE period), relative day is calculated as start (stop) date minus first open-label dose date + 1. The relative day in this situation should be preceded by a 'o'.
- For subjects entering the open-label CZP period (including SFE period), if the start (stop) date occurred after the last dose of open-label CZP drug, the relative day to the most recent open-label CZP dose is calculated as start (stop) date minus most recent open-label CZP dose date. The relative day in this situation should be preceded by a 'o+'.
- 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 'd-'.

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

**Change #16**

**Section 3.5.5:**

The Per-Protocol Set (PPS) will consist of subjects in the FAS without any important protocol deviations that may influence the validity of the data for the primary efficacy variable. Treatment

compliance as defined in Section 7 may also be utilized. Important protocol deviations will be predefined and evaluated at the data evaluation meeting prior to study unblinding database lock.

**Has been changed to the following and Section 3.5.6 added:**

The Per Protocol Set (PPS) will consist of subjects in the FAS without any important protocol deviations that may influence the validity of the data for the primary efficacy variable. Treatment compliance as defined in Section 7 may also be utilized. Important protocol deviations will be predefined and evaluated at the DEM prior to study unblinding at the Week 52 interim database lock. Protocol deviations occurring after Week 52 (for example, for subjects continuing in the SFE period) will not be considered for PPS impact as they occurred after assessment of the primary variable was performed.

**3.5.6 SFE Safety Set**

The SFE Safety Set (SFE-SS) will be defined as all subjects who continued into the SFE Period and who received at least 1 dose of CZP in the SFE Period. Safety data recorded during the SFE will be presented alongside the Week 52 analyses and hence it is not expected that this analysis set will be required.

**Change #17**

**Section 3.6:**

In general, for by visit data collected during the open-label period, subjects will be presented according to the treatments received in the double-blind period (CZP 200mg Q2W and PBO).

Concomitant medications will be summarized according to whether they were taken: during the double-blind period by CZP 200mg Q2W or PBO, or during the open-label period by PBO->OL CZP (patients who were randomized to PBO and then received open-label CZP), CZP->OL CZP (patients who were randomized to CZP and then received open-label CZP), or by the open-label other treatments group. In addition, concomitant medications taken during any CZP medication period (Total CZP) and all concomitant medications (All subjects) will also be summarized. See Section 6.4 for more detail.

AEs will be assigned to periods based on whether they started during the double-blind, open-label CZP or open-label other treatment. AEs will be summarized for the double-blind period by CZP 200mg Q2W or PBO, for the open-label CZP period by PBO->OL CZP or CZP->OL CZP (as above), by the open-label other treatments and by Total CZP (includes randomized CZP 200mg Q2W and AEs starting in the open-label CZP period). See Section 11.2 for more detail.

**Has been changed to:**

In general, for by visit data collected during the open-label period, subjects will be presented according to the treatments received in the double-blind period (CZP 200mg Q2W and PBO).

For the by visit data collected during the SFE period, subjects will be presented in a single CZP 200mg Q2W treatment group.

Concomitant medications will be summarized according to whether they were taken: during the double-blind period by CZP 200mg Q2W or PBO, or during the open-label period by PBO->OL CZP (subjects who were randomized to PBO and then received open-label CZP), CZP->OL CZP (subjects who were randomized to CZP and then received open-label CZP), or by

the open-label other treatments group. In addition, concomitant medications taken during any CZP medication period (Total CZP) and all concomitant medications (All subjects) will also be summarized. See [Section 6.4](#) for more detail.

AEs will be assigned to periods based on whether they started during the double-blind, open-label CZP, open-label other treatment, or SFE periods. AEs will be summarized as follows:

- Double-blind period by CZP 200mg Q2W or PBO,
- Open-label CZP period by PBO->OL CZP or CZP->OL CZP (as above),
- Open-label other treatment period,
- SFE period by OL CZP
- Total CZP (includes randomized CZP 200mg Q2W, AEs starting in the open-label CZP period and AEs starting in the SFE period).

See [Section 11.2](#) for more detail.

#### **Change #18**

##### **Section 3.9:**

Not applicable

##### **Has been changed to:**

The protocol stated that ADAb concentrations above 2.4 unit/ml would be defined as ADAb positive. However, during the trial it was noted that the background sample noise could only be determined during the sample analyses and hence the cut point cannot be pre-specified. Therefore the immunogenicity sections of the SAP were updated to define ADAb positive as being greater than a defined cut point. The cut point used will be agreed after the sample analysis at the DEM prior to database lock and unblinding.

#### **Change #19**

##### **Section 4.3:**

No interim analysis is planned for this study.

##### **Has been changed to:**

An interim analysis is planned after the completion of the double-blind period of the last subject at Week 52. At this time, the database from the double-blind period will be locked, the treatment codes will be made available to relevant to the study reporting team and an interim study report will be written. The investigators and subjects will remain blind to the assigned CZP dose regimen of the double-blind period. After the completion of the SFE period of the last subject, the database will be locked, and a final study report will be written. From the Week 52 visit onward, subjects will be treated with open-label CZP until the last dosing visit of the study (SFE Week 104).

#### **Change #20**

##### **Section 4.2.2:**

- Last Observation Carried Forward (LOCF): If there is missing data at the double-blind time point of interest, the last available post-Baseline measurement will be carried forward to that time point (including any previous scheduled or unscheduled visits). Observations will be carried forward to all missing visits up until the time point of interest even if the subject has withdrawn from the study or entered the open-label period. This will not apply for the analysis of SI joint SPARCC score at Week 12 since there are no measurements collected between Baseline and Week 12.

**Has been changed to:**

- Last Observation Carried Forward (LOCF): If there is missing data at the double-blind time point of interest, the last available post-Baseline measurement will be carried forward to that time point (including any previous scheduled or unscheduled visits in double-blind or open-label periods). Using the date that the observation was recorded, data will be carried forward to all missing double-blind visits up until the time point of interest even if the subject has withdrawn from the study or entered the open-label period. For example, a subject withdrawing from double-blind at Visit 14 /Week 24 and entering open-label CZP will proceed to have a open-label visits at weeks 0, 2 and 4 post double-blind withdrawal followed by assessments every 12 weeks until 52 weeks post randomization. Using the date of the open-label visits, the number of weeks post randomization will be calculated to allow data to be carried forward to all subsequent missing double-blind visits. This will not apply for the analysis of SI joint SPARCC score at Week 12 since there are no measurements collected between Baseline and Week 12.

**Change #21**

**Section 4.2.4:**

The following sentence was added: If time of study treatment and time of event or concomitant medication is available and non partial then it will also be used however if missing or partial, only the dates will be compared.

**Change #22**

**Section 4.3:**

No interim analysis is planned for this study.

**Has been changed to:**

An interim analysis is planned after the completion of the double-blind period of the last subject at Week 52. The timing of the analysis will be when the last subject completes Week 52 visit in accordance with the double-blind, OL-CZP or OL-OT schedule. At this time, all visit based data from the double-blind period up to and including Week 52 visit, all open-label visit based data up to Week 52 visit, all available Follow-up data for subjects already completing the study, all hospitalization/emergency room visit and health care provider consultation data, all study medication discontinuation and all prior and concomitant medication data including concomitant medical procedures, will be locked. Subjects still on the study awaiting Follow-up visit or in SFE, will continue to have AEs, laboratory, vital sign, PK and physical examination data collected per schedule until completion of the Follow-up or SFE period, at which point they will complete the study termination eCRF page. At the time of the interim analysis Week 52 DB

lock, adverse event data entered to date will be cleaned but not locked. This will allow ongoing adverse event entry during the Follow-up and SFE periods. The treatment codes will be made available to the study reporting team and an interim study report will be written. All tables, listings and figures described in this SAP for analysis of the double-blind 52-week treatment period will be produced at the time of the Week 52 DB lock. The Investigators and subjects will remain blind to the assigned CZP dose regimen of the double-blind period.

After the last subject has completed the SFE period, the database will be fully locked, and a final study report will be written. For subjects in the SFE period, from the Week 52 visit onward, subjects will be treated with OL-CZP until Week 156/SFE-WD visit. During this time, only adverse events, concomitant medication related to adverse events, and study drug dispensation information will be collected. Efficacy assessments and study termination data will be collected at Week 156/SFE-WD as described in Table 5-6 of the study protocol. Therefore, at the final DB lock, the following tables and listings will be produced.

- Table 1.1.2: Disposition of Subjects Screened
- Table 1.2.1: Disposition and Discontinuation Reasons
- Table 1.2.2: Number of Subjects Completing each Visit
- Table 1.3: Disposition of Analysis Sets
- Table 1.4.1: Important Protocol Deviations During SFE
- Efficacy tables for analysis of efficacy data collected at Week 156/SFE-WD visit per Table 5-6 of the study protocol
- Table 7.1: Extent of Exposure
- Adverse Event Tables
- Laboratory Tables in order to include any data collected at the Follow-up visit post Week 52 interim database lock for subjects not entering SFE
- Listing 1.2: Subject Disposition
- Listing 1.3: Study Termination
- Listing 1.4: Visit Dates
- Listing 3.1: Important Protocol Deviations
- Listing 4.1: Subject Analysis Sets
- Efficacy listings for all efficacy data collected at Week 156/SFE-WD visit per Table 5-6 of the study protocol
- Adverse Event listings
- Laboratory listings in order to include any data collected at the Follow-up visit post Week 52 interim database lock for subjects not entering SFE.

Regular monitoring of safety data collected during the study 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.

### **Change #23**

#### **Section 4.6:**

Other than the planned analyses based on the PPS, no other efficacy subsets are defined for statistical analyses.

#### **Has been changed to:**

The Randomized Set will be used to evaluate all subjects who were randomized to double-blind study medication. This analysis set will provide additional information on the efficacy analysis by describing findings in the full set of subjects who were randomized and will not exclude those who did not receive at least 1 dose of study medication or did not have a valid Baseline efficacy measurement for ASDAS.

Other than the planned sensitivity analyses based on the PPS and Randomized Set, no other efficacy subsets are defined for statistical analyses.

### **Change #24**

#### **Section 4.8:**

Subgroups for age (<45 and  $\geq$ 45 years), gender (male, female), race (white, other), symptom duration (<5,  $\geq$  5 years), tobacco use (never, current former), HLA-B27 genotype (positive, negative), region (North America, Europe, Asia), prior anti-TNF exposure (yes, no), anti-CZP antibody status (>2.4 units/mL at any post-Baseline assessment,  $\leq$ 2.4 units/mL at all post-Baseline assessments), and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+) will be utilized for summarizing the primary efficacy variable. Only summary statistics will be presented.

#### **Has been changed to:**

Subgroups for age (<45 and  $\geq$ 45 years), gender (male, female), race (white, other), symptom duration (<5,  $\geq$  5 years), tobacco use (never, current former), HLA-B27 genotype (positive, negative), region (North America, Europe, Asia), prior anti-TNF exposure (yes, no), subjects with/without relevant changes to background medication, anti-CZP antibody status (positive, negative as defined in [Section 10](#)), and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+) will be utilized for summarizing the primary efficacy variable. Only summary statistics will be presented.

### **Change #25**

#### **Section 6.2**

- back pain of  $\geq$  3 month duration and age of onset < 45

#### **Has been changed to:**

- back pain of  $\geq$  3 month duration and age of onset < 45 (and per inclusion criteria back pain of  $\geq$  12 month duration and age of onset < 45)

### **Change #26**

#### **Section 7**

---

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 is given as follows:

$$\text{CR} = (\text{Study Duration} - \text{Cumulative Difference}) / \text{Study Duration}$$

The CR ranges between 0 and 1.

To calculate study duration, the date of the Week 50 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 50 visit/last injection date – Baseline date (maximum value is 350 days)

**Has been changed to:**

CR for syringes = # 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). The general formula for the compliance ratio is given as follows:

$$\text{CR for injection day} = (\text{Study Duration} - \text{Cumulative Difference}) / \text{Study Duration}$$

The CR for injection day ranges between 0 and 1.

To calculate double-blind study duration, the date of the Week 50 visit or the last injection date prior to double-blind study treatment discontinuation will be compared to the Baseline date, as shown below:

Double-blind Study Duration (days) = Week 50 visit/last injection date – Baseline date (maximum value is 350 days)

**Change #27**

**Section 8**

All efficacy analyses will be performed using the FAS. The PPS will be used for a sensitivity analysis on the primary endpoint only.

**Has been changed to:**

All efficacy analyses will be performed using the FAS. The PPS and Randomized Set will be used for a sensitivity analysis on the primary endpoint only (using the composite endpoint analysis).

## **Change #28**

### **Section 8.1.3**

The logistic regression analyses of the composite endpoint for ASDAS-MI responder rates at Week 52 (as described in [Section 8.1.2](#)) will also be repeated using the PPS.

#### **Has been changed to:**

The logistic regression analyses of the composite endpoint for ASDAS-MI responder rates at Week 52 (as described in [Section 8.1.2](#)) will also be repeated using the PPS and the Randomized Set.

## **Change #29**

### **Section 8.3.1.10**

The observed data at Week 52 will be used and analyzed according to randomized treatment, regardless of whether or not the subject is still on his/her randomized double-blind study treatment at that time.

#### **Has been changed to:**

The observed data at Week 52 will be used and analyzed according to randomized treatment, regardless of whether or not the subject is still on his/her randomized double-blind study treatment at that time. For example, subjects on open-label CZP or open-label other treatment at the time of their Week 52 visit will have that data included in the analysis including observed data at Week 52.

## **Change #30**

### **Section 8.2.3**

Sensitivity analyses of the secondary binary efficacy variables (ASAS40 at Weeks 12 and 52) as described in [Section 8.2](#) will be as follows:

- Including observed data at Week 52 (or Week 12 as applicable): As described for the primary efficacy variable in [Section 8.1.3.1](#).

#### **Has been changed to:**

Sensitivity analyses of the secondary binary efficacy variables (ASAS40 at Weeks 12 and 52) as described in [Section 8.2](#) will be as follows:

- Including observed data at Week 52 (or Week 12 as applicable): As described for the primary efficacy variable in [Section 8.1.3.1](#). It is possible that a subject completes the double-blind Week 12 visit and withdraws at this visit onto open-label, such that they complete an open-label Week 12 visit in addition. For any visits with data collected at the same visit but in multiple periods, the double-blind visit data will be used in preference to the open-label CZP data which will be used in preference to the open-label other treatment data.

## **Change #31**

### **Section 8.3.1.10**

The SF 36 section was re-written to use the QualityMetric Health Outcomes™ Scoring Software instead of the Ware et al, 2007 scoring algorithm.

### **Change #32**

#### **Section 9.1 paragraph 1**

Certolizumab pegol plasma concentration data will be summarized by treatment group overall, and by antibody status, for each visit at which samples were taken, for the SS.

#### **Has been changed to:**

Certolizumab pegol plasma concentration data will be summarized by treatment group overall, and by Anti-CZP antibody (ADAb) status of individual subjects (as defined below), for each scheduled visit at which samples were taken, for the SS. Samples collected out of window may be excluded from the analysis. The final decision will be determined at the DEM meeting prior to database lock.

### **Change #33**

#### **Section 9.1 paragraph 3**

Individual plasma concentrations for CZP, and anti-CZP levels versus time will be produced on the same graph. The Y-axis on the left will represent the CZP and the Y-axis on the right will show the anti-CZP levels (units/mL).

#### **Has been changed to:**

Individual subject plasma concentrations for CZP, anti-CZP levels, ASDAS scores and dose levels versus time will be produced on the same graph. Two Y-axes on the left will represent the CZP concentrations and anti-CZP levels. The Y-axis on the right will show the ASDAS score. The x-axis will represent time (weeks) with lines showing the treatment and dose.

### **Change #34**

#### **Section 10:**

Frequency tables of anti-CZP antibody (ADAb) status by visit will be presented for the SS.

The number and percentage of subjects with anti-CZP antibody concentrations above 2.4 units/mL will be reported as follows:

- Number and percentage of subjects with ADAb >2.4 units/mL at the time of each visit
- Number and percentage of subjects with ADAb >2.4 units/mL at any visit during treatment (not including post treatment withdrawal or Follow-up visits)
- Number and percentage of subjects with ADAb >2.4 units/mL at any visit including post treatment withdrawal and Follow-up visits.

For the subgroup of subjects with at least 1 anti-CZP antibody level above 2.4 units/mL, the time point of occurrence of the first finding will also be displayed.

#### **Has been changed to:**

Anti-CZP antibody (ADAb) status will be determined for each visit where samples were taken. A cut point used to define ADAb negative and positive will be agreed after the sample analysis at the data evaluation meeting (DEM) prior to database lock and unblinding.

In additional, ADAb status of individual subjects will be determined and summarized using the same cut point:

- ADAb positive is defined as having a value  $>$  a defined cut point while on CZP treatment
- ADAb negative is defined as having no values  $>$  a defined cut point while on CZP treatment

The above definitions will be used for subgroup analyses by anti-CZP antibody status. Note that “while on CZP treatment” in the previous definitions refers to measurements taken when the subject is exposed to CZP study treatment. Measurements at Baseline, at the Follow-up, and at any visit  $>70$  days after last CZP dose are not included in this algorithm.

Note that for purposes of classifying a subject as negative or positive for anti-CZP antibodies, the value must be above the defined cut point and must occur after the subject has had at least one injection of CZP. By definition, a subject cannot be positive for anti-CZP antibodies if they have only received PBO, regardless of how high the value may be. However, a subject randomized to PBO and escaping to CZP may become positive based on values during the OL period while receiving CZP.

A frequency table of anti-CZP antibody status by visit will be presented. In addition, the first occurrence of anti-CZP antibody positive will be tabulated by visit including the cumulative count. That is, each anti-CZP antibody positive subject will be counted only once at the visit where anti-CZP antibody positivity was first observed. The table will also include a summary of the total number of anti-CZP antibody and the total number of positive results observed on treatment (at any visit during treatment not including post treatment withdrawal or Follow-up visits) and at any visit including post treatment withdrawal and Follow-up visits.

### **Change #35**

#### **Section 11.1:**

Exposure to the open-label CZP treatment will also be calculated using the above 4 methods considering only dosing during the time the patient is receiving open-label CZP study medication. Exposure to open-label other treatments will not be calculated.

For the first 3 approaches, tables will summarize exposure days for the placebo, CZP 200mg Q2W and open-label CZP treatment groups.

#### **Has been changed to:**

Exposure to the open-label CZP treatment will also be calculated using the above 4 methods considering only dosing during the time the subject is receiving open-label CZP study medication prior to Week 52 visit. Exposure during the SFE period will be calculated using methods 2, 3 and 4 assuming first dose in the SFE period starts when the subject takes their loading dose/first dose of open-label CZP in the SFE period and ends when the subject has their last administration of study medication (or if missing, the date of withdrawal). Total CZP exposure will be calculated using all 4 methods, by summing the exposure to CZP during the double-blind, open-label CZP and SFE periods. Note that for method 1, the SFE period will be

excluded from the Total CZP as data is not collected on dose administration. Exposure to open-label other treatments will not be calculated.

For the first 3 approaches, tables will summarize exposure for the placebo, CZP 200mg Q2W, open-label CZP treatment groups (split by the treatment taken during the double-blind period), SFE open-label CZP and Total CZP.

**Change #36**

**Section 11.2**

AE start dates will be examined to determine if they are treatment-emergent and attributable to the double-blind treatment period or open-label period. AEs occurring either before double-blind study drug administration (pre-treatment) or outside of the windows described below will be defined as non-treatment emergent.

<b>Table 11-1: Assignment of Adverse Events to Study Periods</b>		
<b>Type of Subject</b>	<b>Assignment Rule Based on AE Start Date</b>	<b>Assignment</b>
Subjects who participate only in the double-blind period	Start date is on or after first study drug administration during the double-blind period and before the last study drug administration + 70 days.	Double-blind treatment-emergent
Subjects who enter the open-label period prior to Week 52	1. Start date is on or after first study drug administration during the double-blind period and before the start date of open-label CZP or other treatment.	Double-blind treatment-emergent
	2. Start date is on or after the start date of open-label CZP and before the start date of open-label other treatment.	Open-label CZP
	3. Start date is on or after the start date of open-label other treatment.	Open-label other treatment

If AE intensity or relationship is missing, then the given event will be imputed as severe or related, respectively.

The frequency of all AEs including TEAEs will be presented for each treatment group separately by System Organ Class, high level term, and preferred term. The data will be displayed as number of subjects experiencing the AEs, percentage of subjects, and number of AEs. Certain summaries will also include exposure-adjusted incidence rates and event rates. Only TEAEs occurring in the double-blind period and open-label period will be presented.

The following AE summaries will be presented:

- Overview of TEAE
- TEAE

- Serious TEAE
- Non-serious TEAE
- TEAE with fatal outcome
- TEAE leading to permanent withdrawal of study medication
- Severe TEAE (included in TEAE table by intensity)
- Drug-related TEAE (included in TEAE table by relationship)
- Serious TEAE by relationship
- Non-serious TEAE by relationship
- Fatal TEAE by relationship
- Injection reactions (Injection site reactions, systemic injection reactions, acute systemic injection reactions, and delayed systemic injection reactions)
- TEAE of special interest (see below)
- TEAE sorted by incidence of PTs in the CZP 200mg group (SOC and HLT will not be utilized)
- Non-serious TEAE with an incidence of at least 5% in 1 of the 2 randomized groups (HLT will not be utilized)
- Non-serious TEAE with an incidence of at least 5% in 1 of the 2 randomized groups by relationship (HLT will not be utilized)
- TEAE subject numbers

The AEs of interest and the approach for summarizing them are described below:

3. Serious infections, including opportunistic infections. Serious infections will be summarized using the Serious TEAE table. No separate table is planned. In addition, opportunistic infections (including tuberculosis) will be presented in a table using UCB-defined search criteria.
4. Malignancies, including lymphoma. These will be presented in 2 tables using the criteria SMQ = "Malignant or unspecified tumours" and SMQ="Malignancies", respectively.
5. Congestive heart failure. These will be manually identified by the study physician from the TEAE table. No separate table is planned. In addition, major adverse cardiac events (MACE) will be presented in a table using UCB-defined search criteria and will include fatal and serious non-fatal myocardial infarction, cerebrovascular events and congestive heart failure.
6. Demyelinating-like disorders. These will be manually identified by the study physician from the previously described TEAE table. No separate table is planned.
7. Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia. These will be presented in a table using the criteria SMQ = "Haematopoietic cytopenias" in the subset of SAEs.

8. Serious bleeding events. These will be presented in a table using the criteria SMQ = “Haemorrhages” in the subset of SAEs.
9. Lupus and lupus-like illness. These will be manually identified by the study physician from the TEAE table. No separate table is planned.
10. Serious skin reactions (eg, Stevens-Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme). These will be manually identified by the study physician from Serious TEAE table. No separate table is planned.

The tables for 1, 2, 3, 5, and 6 above will include the incidence rate with associated 95% CI, and the exposure adjusted event rate.

Although not an AE of interest, hepatic events will also be summarized. They should be identified using the following 5 SMQs: Cholestasis and jaundice of hepatic origin; Hepatic failure, fibrosis, and cirrhosis and other liver damage-related conditions; Hepatitis, noninfectious; Liver-related investigations, signs and symptoms; and Liver-related coagulation and bleeding disturbances.

Because subjects will be able to adjust background medications during the study, some additional AE tables will be prepared in which events occurring while on certain background medications will be summarized. The list of these supportive tables is as follows:

- TEAEs reported after starting a new biologic, but within 10 weeks of stopping CZP
- TEAEs reported one month after changing NSAID type
- TEAEs reported three months after changing or adding a DMARD
- TEAEs reported within one month of an addition of oral corticosteroids or 50% increase in dose

The purpose of these tables will be to assess whether any observed treatment differences in the AE profile may be due, in part, to these prespecified background medications.

A TEAE table will also be presented for the ‘anti-CZP antibody status’ subgroup. Only the subjects exposed to CZP in the double-blind period will be considered and the double-blind TEAEs will be displayed by the antibody status (negative, positive as defined in section 10). A further positive column will be presented by summarizing the TEAEs occurring after the onset of the positive antibody status. The incidence rate will be presented in addition to the associated 95% CI, and the exposure adjusted event rate.

The TEAE incidence rates will also be adjusted for exposure (as calculated using methods 3 and 4 in [Section 11.1](#)) and reported by 100 patient-exposure years. Two approaches will be applied to adjust for exposure. One will only use the first occurrence of an AE with corresponding exposure (exposure-adjusted incidence rate [EAIR]) and the second approach will use all AEs and the entire exposure (exposure-adjusted event rate [EAER]).

For the EAIR, the first occurrence of an TEAEs for a certain treatment group will be utilized (either placebo or CZP 200mg Q2W for the double-blind period or open-label CZP) and divided by the sum of exposure of all subjects in the respective treatment group, where subjects experiencing the respective AE will be censored at the time of occurrence of the AE. The EAIR will be multiplied by a factor of 100 to give a rate per 100 patient-years.

For the EAER, all TEAEs for a certain treatment group will be utilized (either placebo or CZP 200mg Q2W for the double-blind period or open-label CZP) and divided by the sum of exposure of all subjects in the respective treatment group. The EAER will be multiplied by a factor of 100 to give a rate per 100 patient-years.

The summary of TEAEs, serious TEAEs, and TEAEs leading to permanent withdrawal of study medication will include the EAIR (with corresponding 95% exact CI) and EAER. The confidence interval for the EAIR will be based on Poisson rate CIs. The bounds of the exact 95% CIs are estimated by

$$\text{lower bound} = \frac{1}{2 \cdot t} \cdot \chi^2_{2 \cdot (\text{number of patients with events}), 0.025} \quad \text{and} \quad \text{upper bound} = \frac{1}{2 \cdot t} \cdot \chi^2_{2 \cdot (\text{number of patients with events} + 1), 0.975}$$

$\chi^2_{n,\gamma}$  denotes the chi<sup>2</sup>-distribution with n degrees of freedom and quantile  $\gamma$  and t is the exposure censored at the time of occurrence of the AE.

**Has been changed to:**

AE start dates will be examined to determine if they are treatment-emergent and attributable to the double-blind period, open-label period or SFE period. AEs occurring either before double-blind study drug administration (pre-treatment) or outside of the windows described below will be defined as non-treatment emergent.

**Table 11-1: Assignment of Adverse Events to Study Periods**

Type of Subject	Assignment Rule Based on AE Start Date	Assignment
Subjects who do not enter the open-label period prior to Week 52 and do not enter SFE	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days.	Double-blind TEAE
Subjects who do not enter the open-label period prior to Week 52 but do enter SFE after Week 52	Start date is on or after first study drug administration during the double-blind period and prior to loading dose/first dose administration in SFE	Double-blind TEAE
Subjects who enter the open-label period prior to Week 52	1. Start date is on or after first study drug administration during the double-blind period and before the start date of open-label CZP or other treatment.	Double-blind TEAE

<b>Table 11-1: Assignment of Adverse Events to Study Periods</b>		
<b>Type of Subject</b>	<b>Assignment Rule Based on AE Start Date</b>	<b>Assignment</b>
	2. Start date is on or after the start date of open-label CZP and either before the start date of open-label other treatment (for subjects moving into open-label other treatment), or before the loading dose/first dose administration in SFE (for subjects moving into SFE period) or before the last study drug administration during the Open-label CZP period (prior to Week 52) + 70 days (for subjects who complete the Follow-up visit not entering SFE).	Open-label CZP TEAE
	3. Start date is on or after the start date of open-label other treatment.	Open-label other treatment TEAE
Subjects who enter the open-label SFE period (Week 52 completers only)	Start date is on or after the loading dose/first dose administration in SFE.	SFE TEAE

If AE intensity or relationship is missing, then the given event will be imputed as severe or related, respectively.

The frequency of all AEs including TEAEs will be presented for each treatment group separately by System Organ Class, high level term, and preferred term. The data will be displayed as number of subjects experiencing the AEs, percentage of subjects, and number of AEs. Certain summaries will also include exposure-adjusted incidence rates and event rates. TEAEs will be summarized by treatment and period as described in [Section 3.6](#).

The following AE summaries will be presented:

- Overview of TEAE
- TEAE
- Serious TEAE
- Non-serious TEAE
- TEAE with fatal outcome
- TEAE leading to permanent withdrawal of study medication
- Severe TEAE (included in TEAE table by intensity)
- Drug-related TEAE (included in TEAE table by relationship)
- Serious TEAE by relationship

- Non-serious TEAE by relationship
- Fatal TEAE by relationship
- Injection reactions (Injection site reactions, systemic injection reactions, acute systemic injection reactions, and delayed systemic injection reactions)
- TEAE of special interest (see below)
- TEAE sorted by incidence of PTs in the CZP 200mg group (SOC and higher level term [HLT] will not be utilized)
- Non-serious TEAE with an incidence of at least 5% in 1 of the 2 randomized groups (HLT will not be utilized)
- Non-serious TEAE with an incidence of at least 5% in 1 of the 2 randomized groups by relationship (HLT will not be utilized)
- TEAE subject numbers

The AEs of interest and the approach for summarizing them is briefly described below. Further information is provided in the guidance document (AEs of Interest – Cimzia Program 9February2017).

11. Serious infections, including opportunistic infections. Serious infections will be summarized using the Serious TEAE table. No separate table is planned. In addition, opportunistic infections (including tuberculosis) will be presented in a table using UCB-defined search criteria.
12. Malignancies, including lymphoma. These will be presented in 2 tables using the Standardized MedDRA Query (SMQ) criteria “Malignant or unspecified tumours” and “Malignant tumours” respectively.
13. Serious cardiovascular events. These will be presented in a table including all serious TEAEs with the SMQs of “Haemorrhagic central nervous system vascular conditions” and “Ischaemic central nervous system vascular conditions”, including all serious TEAEs with the HLT of “Ischaemic coronary artery disorders” except events coding to PTs of “Chest Pain” or “Chest discomfort” and including all serious TEAEs with HLTs of “Heart Failures NEC”, “Left Ventricular Failures”, or “Right Ventricular Failures” and which also code to the SOC of “Cardiac Disorders”.
14. Congestive heart failure. These will be manually identified by the study physician from the TEAE table. No separate table is planned.
15. Demyelinating-like disorders. These will be manually identified by the study physician from the previously described TEAE table. No separate table is planned.
16. Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia. These will be presented in a table using the criteria SMQ = “Haematopoietic cytopenias” in the subset of Serious TEAEs.
17. Serious bleeding events. These will be presented in a table using the criteria SMQ = “Haemorrhage terms (excl laboratory terms)” in the subset of Serious TEAEs.

18. Lupus and lupus-like illness. These will be manually identified by the study physician from the TEAE table. No separate table is planned.

19. Serious skin reactions (eg, Stevens-Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme). These will be manually identified by the study physician from Serious TEAE table. No separate table is planned.

The tables for 1, 2, 3, 5, and 6 above, will include the exposure-adjusted incidence rate (EAIR) with associated 95% CI, and the exposure-adjusted event rate (EAER) as described below.

Although not officially AEs of interest, the following events will also be summarized as described in the AEs of Interest – Cimzia Program 9 February 2017 guidance document.

20. Hepatic events. These will consist of a subset of all TEAEs, identified using the following 5 SMQs: Cholestasis and jaundice of hepatic origin; Hepatic failure, fibrosis, and cirrhosis and other liver damage-related conditions; Hepatitis, noninfectious; Liver-related investigations, signs and symptoms; and Liver-related coagulation and bleeding disturbances.

21. Hypersensitivity reactions and anaphylactic reactions. Hypersensitivity reactions will be summarized as any TEAEs occurring on the same day or the day after injection was received, which code to the following preferred terms: Administration site hypersensitivity, Documented hypersensitivity to administered product, Drug hypersensitivity, Hypersensitivity, Hypersensitivity vasculitis, Infusion site hypersensitivity, Injection site hypersensitivity, Medical device site, hypersensitivity, Type II hypersensitivity or Type IV hypersensitivity reaction. Anaphylactic reactions will be defined using an algorithmic approach as described in the “AEs of Interest – Cimzia Program 9 February 2017” guidance document.

Because subjects will be able to adjust background medications during the study, some additional AE tables will be prepared in which events occurring while on certain background medications will be summarized. The list of these supportive tables is as follows:

- TEAEs reported after starting a new biologic, but within 10 weeks of stopping CZP
- TEAEs reported one month after changing NSAID type
- TEAEs reported three months after changing or adding a DMARD
- TEAEs reported within one month of an addition of oral corticosteroids or 50% increase in dose

The purpose of these tables will be to assess whether any observed treatment differences in the AE profile may be due, in part, to these prespecified background medications.

A TEAE table will also be presented for the ‘anti-CZP antibody status’ subgroup defined in [Section 10](#). Only subjects exposed to CZP and only TEAEs occurring after first dose of CZP will be considered. A further column will be presented summarizing TEAEs occurring after the onset of the positive antibody status. The EAIR will be presented in addition to the associated 95% CI, and the EAER.

The TEAE incidence rates will also be adjusted for exposure (as calculated using methods 3 and 4 in [Section 11.1](#)) and reported by 100 subject years of exposure. 100 subject-years is defined as the sum of the exposure / number of subjects \*100. Two approaches will be applied to adjust

for exposure. One will only use the first occurrence of an AE with corresponding exposure (EAIR) and the second approach will use all AEs and the entire exposure (EAER).

For the EAIR, the first occurrence of an TEAEs for a certain treatment group will be utilized (either placebo or CZP 200mg Q2W for the double-blind period or open-label CZP) and divided by the sum of exposure of all subjects in the respective treatment group, where subjects experiencing the respective AE will be censored at the time of occurrence of the AE. The EAIR will be multiplied by a factor of 100 to give a rate per 100 subject-years.

For the EAER, all TEAEs for a certain treatment group will be utilized (either placebo or CZP 200mg Q2W for the double-blind period or open-label CZP) and divided by the sum of exposure of all subjects in the respective treatment group. The EAER will be multiplied by a factor of 100 to give a rate per 100 subject-years.

The summary of TEAEs, serious TEAEs, and TEAEs leading to permanent withdrawal of study medication will include the EAIR (with corresponding 95% exact CI) and EAER. The confidence interval for the EAIR will be based on Poisson rate CIs. The bounds of the exact 95% CIs are estimated by

$$\text{lower bound} = \frac{1}{2 \cdot t} \cdot \chi^2_{2, \text{number of patients with events}, 0.025} \quad \text{and} \quad \text{upper bound} = \frac{1}{2 \cdot t} \cdot \chi^2_{2, (\text{number of patients with events} + 1), 0.975}$$

$\chi^2_{n, \gamma}$  denotes the chi<sup>2</sup>-distribution with n degrees of freedom and quantile  $\gamma$  and t is the exposure censored at the time of occurrence of the AE.

### Change #37

#### Section 11.4.1:

The number and percent of subjects with abnormal vital sign values will be summarized by visit and for any visit.

#### Has been changed to:

The number and percent of subjects with abnormal vital sign values will be summarized by visit and for any visit. Abnormal vital sign results are defined as follows.

<b>Table 11-2: Definition of Abnormal Vital Signs</b>		
<b>Vital Sign Parameter</b>	<b>Abnormally Low</b>	<b>Abnormally High</b>
Systolic Blood Pressure (mmHg)	<= 90 and decrease of >=20	>=180 and increase of >=20
Diastolic Blood Pressure (mmHg)	<=50 and decrease of >=15	>=105 and increase of >=15
Pulse Rate (bpm)	<=50 and decrease of >=15	>=120 and increase of >=15

## **Change #38**

### **Section 11.4.3.3:**

Physical examination will be performed at Screening, Baseline, Weeks 2, 4, 8, 12, 16, 24, 36, 52/WD and at the Follow-up visit. Physical examination findings will be recorded in the CRF only at Screening. Physical examination data will be listed only.

### **Has been changed to:**

Physical examination will be performed at Screening, Baseline, Weeks 2, 4, 8, 12, 16, 24, 36, 52/WD and at the Follow-up visit (8 weeks after the Week 52/WD visit for subjects not participating in the SFE period). Physical examination findings will be recorded in the eCRF only at Screening. Physical examination data will be listed only.

## **14.3 AMENDMENT 3**

### **Rationale for the amendment**

The analysis plan was updated to ensure consistency with protocol amendment 4, which included the addition of ASQoL, Nocturnal spinal pain and Uveitis as secondary endpoints and an alternative primary endpoint and testing hierarchy for Canada (and any other country where applicable or where requested by Regulatory Authorities). All modifications are detailed below.

### **Change #1**

General: Minor textual clarifications throughout about data handling for the SFE period.

### **Change #2**

Sections 2.2, 4.2.1, 8, 8.1.3, 8.2.2, 8.2.2.1, 8.2.2.2, 8.2.3, 8.2.3.1, 8.2.3.2, 8.3.1.2 updated to correspond with protocol amendment 4. The following variables were updated:

- ASAS40 was added as the primary endpoint for Canada (and any other country where applicable or where requested by Regulatory Authorities), hence an alternative set of secondary endpoints and sequential testing hierarchy were provided.
- Secondary endpoint added: Change from Baseline in ASQoL at Week 52
- Secondary endpoint added: Change from Baseline in ASQoL at timepoints other than Week 52
- Secondary endpoint added: Change from Baseline in nocturnal spinal pain (NRS) at Week 52
- Secondary endpoint added: Number of subjects with Anterior Uveitis (AU) or new AU flares through Week 52

### **Change #3**

Section: 2.2.3:

Anti-CZP antibody (ADAb) concentrations will be measured and summarized at Baseline, Weeks 1, 2, 4, 12, 24, 36, 52/WD, and the Follow-up visit (if applicable). In addition, the anti-CZP Ab concentrations will be assessed in subjects who withdraw from study drug and

transition to open-label CZP. The number and percentage of subjects with anti-CZP antibody concentrations above a defined cut point will be reported as follows:

- Number and percentage of subjects with ADA<sub>b</sub> > defined cut point at the time of each visit
- Number and percentage of subjects with ADA<sub>b</sub> > defined cut point at any visit during treatment (not including post treatment withdrawal or Follow-up visit)
- Number and percentage of subjects with ADA<sub>b</sub> > defined cut point at any visit including post treatment withdrawal or Follow-up visit

The cut point to use will be agreed after the sample analysis at the data evaluation meeting (DEM) prior to database lock and unblinding.

Was changed to:

Anti-drug antibody (ADA<sub>b</sub>) (Anti-CZP antibody) concentrations will be assessed at Baseline and Weeks 1, 2, 4, 12, 24, 36, 52/WD, and the FU visit (8 weeks after the Week 52/WD visit). In addition, the ADA<sub>b</sub> will be assessed in subjects who withdraw from double-blind study drug and transition to open-label CZP prior to Week 52.

Determination of ADA<sub>b</sub> will be done using a validated screening, confirmation and titration ADA<sub>b</sub> bridging assay, with potential further characterization by a neutralizing antibody assay. The immunogenicity data will be processed according to dedicated Bioanalytical Analysis plans.

The following variables will be analyzed as described in Section 10:

- Anti-CZP antibodies at Baseline, Weeks 1, 2, 3, 12, 24, 36 and 52/WD
- Status of ADA<sub>b</sub> (including overall, baseline and treatment-emergent classification)
- ADA<sub>b</sub> response characterized for their neutralizing potential

#### **Change #4**

Section 2.2.4.1 and 11.2: Summaries of the following were updated:

- Injection reaction summaries were removed.
- TEAEs with CZP incidence  $\geq 1\%$  and CZP incidence > PBO incidence was added.
- Serious Cardiovascular and demyelinating-like disorder were updated to align to latest UCB standards.
- The following was changed to be optional: TEAEs reported after: new biologic, changing NSAID, changing DMARD or addition of oral corticosteroid.

#### **Change #5**

Section 3.1: The following paragraph was removed:

If imputation is being performed, which by definition results in all subjects having available data for analysis, then summary statistics will not present the 'n' (number of available measurements).

### **Change #6**

Section 3.2.1.2, 3.2.1.3, 3.2.1.4, 3.2.1.5: Clarifications were made to the period definitions to ensure a clearer definition of the start and end of each period.

### **Change #7**

Section 3.5.4: The Full Analysis Set (FAS) will consist of all subjects in the RS who have received at least 1 dose of study medication and have a valid Baseline efficacy measurement for ASDAS.

Was changed to:

The Full Analysis Set (FAS) will consist of all subjects in the RS who have received at least 1 dose of study medication.

### **Change #8**

Section 3.6 and 11.2: were updated to provide clarity on how Adverse events are assigned to study periods. An additional Total CZP AE summary was added that includes (includes randomized CZP 200mg Q2W and AEs starting in the open-label CZP period).

### **Change #9**

Section 4.8 and 8.1.3: Baseline SPARCC scores ( $<5$  and  $\geq 5$ ) was added as a subgroup analyses. Anti-CZP antibody was removed and replaced with efficacy-antibody analysis in section 10. Data handling for subjects missing the 2<sup>nd</sup> screening CRP value was added.

### **Change #10**

Section 6.1: Age groups were updated

### **Change #11**

Section 6.4 was updated to clarify how concomitant medications are assigned to periods

### **Change #12**

Section 7: Details for the derivation of Total expected syringes was added and the ratio of compliance categories were extended to present  $>1$ - $<1.2$  and  $>1.2$ .

### **Change #13**

Section 8.1.1 and 8.1.3: Derivation of ASDAS-MI was updated to include handling of subjects with lowest score possible (0.636) post-baseline to include them as responders. A sensitivity analysis was added using the definition of improvement  $\geq 2.0$  relative to Baseline alone.

### **Change #14**

Section 8.1.3.2: Clarified that imputations would be limited to in range results and the MCMC would include measurements from the previous visit as explanatory variables in the model.

### **Change #15**

Section 8.1.3.3: Tipping point deltas updated to vary for placebo and CZP subjects.

### **Change #16**

Section 8.1.3.3: Due to ASAS40 now bring a primary analysis (for Canada), sensitivity analysis were added using methods similar to ASDAS-MI.

### **Change #17**

Section 8.2.1.6 was added since ASQoL is now a secondary endpoint. The text was updated from: 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. If three 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 three items are missing, the total score will be left missing.

To: The ASQoL, a validated disease-specific 18-item questionnaire, has been developed specifically for measuring HRQoL in subjects with AS (Doward et al, 2003). An nr-axSpA specific scoring algorithm is presently being developed, but it is unclear when this algorithm will become available relative to the timing of the AS0006 database lock. Should this alternative scoring approach become available prior to finalization of the clinical study report, analyses of ASQoL data may be repeated using this alternate scoring approach appropriate for nr-axSpA population and methods will be described either in a SAP amendment or within the Documentation of Statistical Methods appendix to the study report.

### **Change #18**

Section 8.2.1.7: Section added and text moved from Other endpoints for Nocturnal Spinal Pain since promotion to secondary endpoint.

### **Change #19**

Section 8.2.2.1: The following text was added since Uveitis was promoted to a secondary endpoint. The number of subjects with new onset post-Baseline AU or new AU flares through Week 52 is a categorical secondary efficacy endpoint for all regions. Subjects will be classified as having post-Baseline Uveitis: a TEAE of Uveitis, a new diagnosis of Uveitis or a flare of uveitis at any visit post-Baseline which is on or before the Week 52 visit. All patients without evidence of uveitis events, new diagnosis or flares will be considered not to have post-Baseline Uveitis. As a dichotomous outcome variable, treatment groups will be compared for differences using logistic regression including treatment, region and MRI/CRP classification as explanatory variables. Subjects with no evidence of new AU flares will be considered as having no flare.

### **Change #20**

Section 8.2.2.2: Step by step description of referenced-based multiple imputation was updated to allow use of the SAS® MNAR option in Proc MI available in version 9.4.

### **Change #21**

Section 8.3: List extended to include EQ-5D which was missed in error. WPS nonparametric bootstrap-t method replaced with the more conventional MMRM analysis.

Section 8.3.1.12: LOCF removed for WPS analysis

Appendix 13.1: Non-parametric bootstrap-t method was removed.

---

**Change #22**

Section 8.3.1.11: ASAS-NSAID analysis updated to summarize by periods using the start/end of concomitant medication dates. The following paragraph was added:

For the observed case analysis, only intervals that end on or prior to end of the double-blind treatment period will be included. For the LOCF analysis, all data will be included for all intervals up to Week 52 (1 year) irrespective of whether the subject is on double-blind or open-label study medication.

**Change #23**

Section 8.3.1.13: MOS sleep scale analysis updated to use QualityMetric Health Outcome scoring software.

**Change #24**

Section 9.1: The following paragraph was inserted to replace evaluation of evaluable PK data at the DEM meeting prior to DB lock. All CZP concentrations will be reviewed after unblinding by study statistician and clinical pharmacologist to ensure all values are plausible, and values deemed anomalous will be excluded from summary statistics.

**Change #25**

Section: 10: Immunogenicity analysis section was replaced with latest best practices.

**Change #26**

Section: 11.1: A Total CZP exposure summary was added that includes (includes randomized CZP 200mg Q2W and open-label CZP period but excludes SFE CZP).

**Change #27**

Section: 11.3: A shift table of liver function test change from baseline to maximum post-baseline category was added.

**Change #28**

Section: 11.4.3.4: A table of QuantiFERON-TB Gold Tuberculosis Test Results was added.

## 14.4 AMENDMENT 4

### Rationale for the amendment

The analysis plan was updated to include the additional analyses required following the changes to the protocol outlined in protocol amendment 5. All modifications are detailed below.

#### Change #1

Section 1 was changed from: This SAP describes the analysis of the double-blind 52-week treatment period and the 10 week follow up period of the C-axSpAnd (AS0006) study. It is designed to support a Clinical Study Report (CSR), is compliant with International Conference on Harmonization (ICH) guidelines and is based on the Protocol Amendment 3 dated 13 February 2017.

To: This SAP describes the analysis of the double-blind 52-week treatment period and the 2 year open-label CZP extension of the C-axSpAnd (AS0006) study. It is designed to support a Clinical Study Report (CSR), is compliant with International Conference on Harmonization (ICH) guidelines and is based on the Protocol Amendment 5 dated 17 December 2018.

#### Change #2

Section 2.2.1.3: The following variables will be analyzed at scheduled time points through Week 52:

- Change from BASMI linear
- ASDAS disease activity (ASDAS-ID, ASDAS-MD, ASDAS-HD, ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI)
- Change from Baseline in SI joint SPARCC score at Week 52 and ASspiMRI-a in the Berlin modification at Week 12 and Week 52
- Proportion of subjects with SI joint SPARCC score  $<2$  at Week 12 and Week 52

Was changed to: The following variables will be analyzed at scheduled time points through Week 52 and at Week 156 in accordance with the protocol schedule of assessments:

- Change from Baseline in BASMI linear
- ASDAS disease activity (ASDAS-ID, ASDAS-LD/ASDAS-MD, ASDAS-HD, ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI)
- Change from Baseline in SI joint SPARCC score at Week 52 and Week 156 and ASspiMRI-a in the Berlin modification at Week 12 and Week 52
- Proportion of subjects with SI joint SPARCC score  $<2$  at Week 12, Week 52, and Week 156

### Change #3

Section 2.2.4: Additional sub-headings (2.2.4.1 Secondary safety variables and 2.2.4.2 Other safety variables) were added with the previous headings for Adverse Events becoming 2.2.4.1.1, Laboratory parameters becoming 2.2.4.2.1, vital signs becoming 2.2.4.2.2 and other safety variables becoming 2.2.4.2.3; in order to separate out the secondary and other safety variables to be in line with the updated protocol.

### Change #4

Section 2.3.1: Starting at Week 64, all subjects are to visit the site approximately every 12 weeks for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156.

Was changed to: Starting at Week 64, all subjects are to visit the site approximately every 12 weeks ( $\pm 2$  weeks) for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156/Safety Follow-up Extension-Early Withdrawal Visit (SFE-WD) ( $\pm 1$  week).

### Change #5

Section 2.3.2: The end of the study will be defined as the date of the last subject's last visit, defined as the Follow-up Visit 8 weeks after Week 52/WD visit for subjects not participating in the SFE Period, or the last visits of the SFE Period.

Was changed to: The end of the study will be defined as the date of the last subject's last visit, defined as the Follow-up Visit 8 weeks after Week 52/WD visit for subjects not participating in the SFE Period, or the last visit of the SFE Period (Week 156/SFE-WD).

### Change #6

Section 3.2.1.5: The SFE period ends when the subject discontinues from or completes the study following SFE enrollment.

Was changed to: The SFE period ends when the subject discontinues from or completes the study following SFE enrollment (Week 156/SFE-WD).

### Change #7

[Section 3.5.7](#) SFE Full Analysis Set was added, defined as: The SFE Full Analysis Set (SFE-FAS) will be defined as all subjects in the FAS, who continued into the SFE Period and who received at least 1 dose of CZP in the SFE Period. This analysis set will be used to analyze efficacy data at Week 156/SFE-WD.

### Change #8

[Section 4.2.3](#) Other efficacy variables: Analyses of other binary efficacy variables will be treated in the same way as the primary analysis of the primary efficacy variable. That is, a subject will be considered a responder only if the response is achieved and if the subject is still on their randomized double-blind study treatment at the time when the variable is evaluated. Analyses of other continuous efficacy variables will be based on LOCF. A sensitivity analyses will be performed on all other efficacy variables using observed case analysis as described for the primary endpoint in [Section 4.2.1](#).

Was changed to:

For the analysis of the double-blind treatment period, other binary efficacy variables will be treated in the same way as the primary analysis of the primary efficacy variable (composite endpoint). That is, a subject will be considered a responder only if the response is achieved and if the subject is still on their randomized double-blind study treatment at the time when the variable is evaluated. Analyses of other continuous efficacy variables will be based on LOCF. A sensitivity analyses will be performed on all other efficacy variables using observed case analysis as described for the primary endpoint in Section 4.2.1.

For analysis of the Week 156/SFE-WD visit data, binary efficacy variables will be treated in the same way as the primary analysis of the primary efficacy variable (Composite Endpoint). That is, a subject will be considered a responder only if the response is achieved and the last administration of OL-CZP study medication is on or after Week 154/Day 1079 (post randomization). Subjects with missing data at Week 156/SFE-WD; or who have withdrawn treatment prior to Day 1079; or those without response at Week 156/SFE-WD, are counted as non-responders. Hence all patients in the SFE-FAS analysis set will be included in this analysis.

Two observed case sensitivity analyses of binary efficacy variables at Week 156/SFE-WD will be performed. The Observed Case analysis at Week 156, will consider a subject a responder only if the response is achieved at the Week 156 visit and the subject completed the study. Non-response at Week 156 visit will be counted as non-responders. Early withdrawals or subjects with missing data will be considered missing in this analysis. The Observed Case analysis at Week 156/SFE-WD will consider a subject as a responder if response is achieved at Week 156/SFE-WD visit such that both subjects with a Week 156 visit (subjects who complete) and subjects with early withdrawal visits and efficacy results meeting the criteria of response, will be counted as a responder. Subjects with missing data will be considered missing in this analysis.

For continuous efficacy variables, the observed data at Week 156/SFE-WD visit will be analyzed. This will summarize efficacy data for completers and subjects with early withdrawal visits during SFE together. In addition, the efficacy data collected at Week 156 for completers only will be summarized.

#### **Change #9**

**Section 4.2.5** Other considerations: Descriptive summaries for all efficacy variables based on observed case data (including only subjects still on their randomized double-blind study treatment) will also be prepared.

Was changed to: Descriptive summaries for all efficacy variables collected during the double-blind treatment period based on observed case data (including only subjects still on their randomized double-blind study treatment) will also be prepared.

#### **Change #10**

**Section 4.3** Interim analyses and data monitoring:

The treatment codes will be made available to the study reporting team and an interim study report will be written. All tables, listings and figures described in this SAP will be produced at the time of the Week 52 DB lock. The Investigators and subjects will remain blind to the assigned CZP dose regimen of the double-blind period.

After the last subject has completed the SFE period, the database will be fully locked, and a final study report will be written. For subjects in the SFE period, from the Week 52 visit onward, subjects will be treated with OL-CZP until the last dosing visit of the study (SFE Week 104). During this time, only adverse events, concomitant medication related to adverse events, and study drug dispensation information will be collected, in addition the study termination data at completion or early withdrawal. Therefore, at the final DB lock, only the following tables and listings will be produced.

- Table 1.1.2: Disposition of Subjects Screened
- Table 1.2.2: Number of Subjects Completing each Visit
- Table 1.3: Disposition of Analysis Sets
- Table 1.4: Important Protocol Deviations
- Table 7.1: Extent of Exposure
- All tables in section 8 (adverse events, 31 tables in total with the exception of tables counting AEs reported after starting a new biologic, changing NSAIDs, changing DMARDs, or adding oral corticosteroids).
- Listing 1.2: Subject Disposition
- Listing 1.3: Study Termination
- Listing 1.4: Visit Dates
- Listing 3.1: Important Protocol Deviations
- Listing 4.1: Subject Analysis Sets
- All listings in section 7 (adverse events, 6 listings in total)

Note that if a substantial number of subjects are still awaiting their Follow-up visit when the interim analysis Week 52 DB lock occurs (and are therefore not entering the SFE), vital sign, laboratory, PK and physical examination outputs may be updated prior to the SFE analysis to include the information collected at the Follow-up visit.

Was changed to:

The treatment codes will be made available to the study reporting team and an interim study report will be written. All tables, listings and figures described in this SAP for analysis of the double-blind treatment period will be produced at the time of the Week 52 DB lock. The Investigators and subjects will remain blind to the assigned CZP dose regimen of the double-blind period.

After the last subject has completed the SFE period, the database will be fully locked, and a final study report will be written. For subjects in the SFE period, from the Week 52 visit onward, subjects will be treated with OL-CZP until Week 156/SFE-WD visit. During this time, adverse events, concomitant medication related to adverse events, and study drug dispensation information will be collected in an ongoing way, in addition to efficacy assessments and study termination data at Week 156/SFE-WD as described in Table 5-6 of the study protocol. Therefore, at the final DB lock, only the following tables and listings will be produced.

- Table 1.1.2: Disposition of Subjects Screened
- Table 1.2.1: Disposition and Discontinuation Reasons
- Table 1.2.2: Number of Subjects Completing each Visit
- Table 1.3: Disposition of Analysis Sets
- Table 1.4.1: Important Protocol Deviations During SFE
- Efficacy tables for analysis of efficacy data collected at Week 156/SFE-WD visit per Table 5-6 of the study protocol
- Table 7.1: Extent of Exposure
- Adverse Event Tables
- Laboratory Tables in order to include any data collected at the Follow-up visit post Week 52 interim database lock for subjects not entering SFE
- Listing 1.2: Subject Disposition
- Listing 1.3: Study Termination
- Listing 1.4: Visit Dates
- Listing 3.1: Important Protocol Deviations
- Listing 4.1: Subject Analysis Sets
- Efficacy listings for all efficacy data collected at Week 156/SFE-WD visit per Table 5-6 of the study protocol
- Adverse Event listings
- Laboratory listings in order to include any data collected at the Follow-up visit post Week 52 interim database lock for subjects not entering SFE.

#### **Change #11**

**Section 4.6:** Other than the planned sensitivity analyses based on the PPS and Randomized Set, no other efficacy subsets are defined for statistical analyses.

Was changed to: Other than the planned sensitivity analyses based on the PPS and Randomized Set, no other efficacy subsets are defined for statistical analyses during the double-blind study period. The SFE-FAS is defined for analysis of efficacy data at the Week 156/SFE-WD visit.

#### **Change #12**

**Section 8:** All efficacy analyses will be performed using the FAS. The PPS and Randomized Set will be used for a sensitivity analysis on the primary endpoint only (using the composite endpoint analysis).

Was changed to: All efficacy analyses will be performed using the FAS. The PPS and Randomized Set will be used for a sensitivity analysis on the primary endpoint only (using the composite endpoint analysis). The SFE-FAS will be used for analysis of efficacy data at the Week 156/SFE-WD visit.

### Change #13

Section 8.1.1: Disease activity categories based on ASDAS are as follows:

- ASDAS-Inactive Disease (ASDAS-ID): ASDAS <1.3
- ASDAS-Moderate Disease (ASDAS-MD): ASDAS  $\geq$ 1.3, <2.1
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS  $\geq$ 2.1,  $\leq$ 3.5
- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

Was changed to: Disease activity categories based on ASDAS are as follows:

- ASDAS-Inactive Disease (ASDAS-ID): ASDAS <1.3
- ASDAS-Moderate Disease (ASDAS-MD), ASDAS-Low Disease (ASDAS-LD): ASDAS  $\geq$ 1.3, <2.1
- ASDAS-High Disease Activity (ASDAS-HD): ASDAS  $\geq$ 2.1,  $\leq$ 3.5
- ASDAS-very High Disease Activity (ASDAS-vHD): ASDAS >3.5

Note that ASDAS-MD (used in study protocol amendments 4 and earlier) is identical to the definition of ASDAS-LD used in the study protocol amendment 5. ASDAS-MD is used throughout this SAP and in the analysis outputs in order to retain consistency with the Week 52 interim database lock analyses.

### Change #14

Section 8.3 title was changed from “Analysis of other efficacy variables” to “Analysis of other efficacy variables during the double-blind treatment period”. An additional section to describe the SFE analysis was added as shown below: Section 8.4 Analysis of other efficacy variables during SFE

In order to assess long-term efficacy of CZP for subjects with nr-axSPA, the following summaries will be presented on the SFE-FAS:

- The number and percentage of patients achieving response for all categorical variables collected at Week 156/SFE-WD (ASDAS-MI reduction of  $\geq$ 2.0 or lowest score possible, ASDAS-MI reduction of  $\geq$ 2.0, ASDAS-CII, ASAS40, ASAS20, ASAS partial remission, BASDAI 50 and SPARCC score < 2). In addition, a 2-sided exact (Clopper-Pearson) 95% confidence interval for the percentage of patients achieving response will be presented. As described in Section 4.2.3, the summary will include a composite endpoint analysis, observed case analysis at Week 156 and observed case analysis at Week 156/SFE-WD.
- Subgroup analysis of ASDAS-MI responders (reduction of  $\geq$ 2.0 or lowest score possible) and ASAS40 responders by Age group, Gender, Symptom Duration, HLA-B27 Genotype and MRI/CRP classification (as defined in section 4.8) using the composite endpoint analysis described in Section 4.2.3 including 2-sided exact (Clopper-Pearson) 95% confidence intervals for the percentage of patients achieving response.
- The number and percentage of patients with ASDAS-ID, ASDAS-MD, ASDAS-HD, ASDAS-vHD at Week 156 (completers only) and at Week 156/SFE-WD

- Summary statistics along with the changes from baseline for ASDAS, BASDAI (total score, individual questions 1, 2, 3, 4 and morning stiffness [Q5, Q6 average]), BASFI, SI Joint SPARCC, PGADA, Nocturnal spinal pain, total spine pain, CRP and ASQoL, collected at Week 156/SFE-WD. As described in Section 4.2.3, observed data for completers only (at Week 156) in addition to observed data for all subjects (completers and subjects with early withdrawal visits) will be presented.

## 14.5 AMENDMENT 5

### Rationale for the amendment

The analysis plan was mainly updated to include the additional imputation rules, add an additional table and to change the derivation of the start of the SFE period. All modifications are detailed below.

### Change #1

The list of abbreviations was changed from:

ADAb	Anti-drug antibody
ADaM	analysis data model
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AP	alkaline phosphatase
ASAS	Assessment of SpondyloArthritis international Society
ASAS20, 40	Assessment of SpondyloArthritis international Society 20%, 40% response criteria
ASAS5/6	Assessment of SpondyloArthritis international Society 20% improvement in 5 of 6 domains
AS	ankylosing spondylitis
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASDAS-CII	Ankylosing Spondylitis Disease Activity Score – Clinically Important Improvement
ASDAS-HD	Ankylosing Spondylitis Disease Activity Score – High Disease activity

---

ASDAS-ID	Ankylosing Spondylitis Disease Activity Score – Inactive Disease
ASDAS-LD	Ankylosing Spondylitis Disease Activity Score – Low Disease
ASDAS-MD	Ankylosing Spondylitis Disease Activity Score – Moderate Disease
ASDAS-MI	Ankylosing Spondylitis Disease Activity Score – Major Improvement
ASDAS-vHD	Ankylosing Spondylitis Disease Activity Score – very High Disease activity
ASQoL	Ankylosing Spondylitis Quality of Life
ASspiMRI-a	Ankylosing Spondylitis spine MRI acuity
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AU	Anterior Uveitis
AxSpA	axial spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASDAI50	Bath Ankylosing Spondylitis Disease Activity Index 50% response criteria
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BMI	body mass index
BMP	bone morphogenic protein
BP	bodily pain
CI	confidence interval
CK	creatinine kinase
CR	compliance ratio
CRP	C-reactive protein

---

CSR	clinical study report
CV	coefficient of variation
CZP	certolizumab pegol
DEM	Data Evaluation Meeting
DKK1	dickkopf-related protein 1
DRL	Drug Reference List
DMARD	disease-modifying antirheumatic drug
EAIR	exposure-adjusted incidence rate
EAER	exposure-adjusted event rate
eCRF	electronic case report form
EQ-5D	EuroQoL Health Status Questionnaire (5 dimensions)
ER	emergency room
ES	Enrolled Set
FAS	Full Analysis Set
GH	general health
HCQ	hydroxychloroquine
HLA-B27	human leukocyte antigen B27
HLT	High Level Term
HRQoL	health-related quality of life
ICH	International Conference on Harmonisation
IXRS	interactive response system
IBD	inflammatory bowel disease
LOCF	last observation carried forward
LLN	lower limit of normal

---

LLOQ	lower limit of quantification
LLT	Low Level Term
MASES	Maastricht Ankylosing Spondylitis Enthesitis Score
MCS	Mental Component Summary
MedDRA <sup>®</sup>	Medical Dictionary of Regulatory Activities <sup>®</sup>
MH	mental health
MMP	matrix metalloproteinase
mNY	Modified New York (criteria)
MOS	Medical Outcomes Study
MRI	magnetic resonance imaging
MTX	methotrexate
nADA	neutralizing ADA <sub>b</sub>
nr-axSpA	nonradiographic axSpA
NRS	Numeric Rating Scale
NRI	non-response imputation
NSAID	nonsteroidal anti-inflammatory drug
OC	observed case
OL	Open-label
OL-CZP	Open-label CZP
OL-OT	Open-label Other Treatment
PBO	placebo
PCS	Physical Component Summary
PF	physical Function
PGADA	Patient's Global Assessment of Disease Activity

---

PhGADA	Physician's Global Assessment of Disease Activity
PPS	Per Protocol Set
PT	Preferred Term
Q	question
Q1	1 <sup>st</sup> Quartile
Q3	3 <sup>rd</sup> Quartile
Q2W	every 2 weeks (every other week)
QoL	quality of life
RBC	red blood cell
RCTC	Rheumatology Common Toxicity Criteria
RE	role emotional
RP	role physical
RS	Randomized Set
SAARD	slow-acting anti-rheumatic drug
SAP	Statistical Analysis Plan
sc	subcutaneous(ly)
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SF	social functioning
SFE	Safety Follow-up Extension
SFE-WD	Safety Follow-up Extension-Early Withdrawal Visit
SFE-FAS	Safety Follow-up Extension Full Analysis Set
SF-36	Short-Form 36-Item Health Survey

---

SI	sacroiliac
SJC	swollen joint count
SM	spinal mobility
SMQ	Standardized MedDRA Query
SOC	System Organ Class
SOP	standard operating procedure
SpA	spondyloarthritis
SPARCC	Spondyloarthritis Research Consortium of Canada
sqrt	square root
SS	Safety Set
SSZ	sulfasalazine
STIR	short-tau-inversion recovery
TB	tuberculosis
TEAE	treatment-emergent adverse event
TJC	tender joint count
TNF	tumor necrosis factor
USA	United States of America
ULN	upper limit of normal
VAS	visual analog scale
VT	vitality
VU	vertebral units
WBC	white blood cell
WHO	World Health Organization
WPS	Work Productivity Survey

To:

ADAb	anti-drug antibody
ADaM	analysis data model
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AP	alkaline phosphatase
ASAS	Assessment of SpondyloArthritis international Society
ASAS20, 40	Assessment of SpondyloArthritis international Society 20%, 40% response criteria
ASAS5/6	Assessment of SpondyloArthritis international Society 20% improvement in 5 of 6 domains
AS	ankylosing spondylitis
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASDAS-CII	Ankylosing Spondylitis Disease Activity Score – Clinically Important Improvement
ASDAS-HD	Ankylosing Spondylitis Disease Activity Score – High Disease activity
ASDAS-ID	Ankylosing Spondylitis Disease Activity Score – Inactive Disease
ASDAS-LD	Ankylosing Spondylitis Disease Activity Score – Low Disease
ASDAS-MD	Ankylosing Spondylitis Disease Activity Score – Moderate Disease
ASDAS-MI	Ankylosing Spondylitis Disease Activity Score – Major Improvement
ASDAS-vHD	Ankylosing Spondylitis Disease Activity Score – very High Disease activity
ASQoL	Ankylosing Spondylitis Quality of Life

---

ASspiMRI-a	Ankylosing Spondylitis spine MRI acuity
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AU	anterior uveitis
AxSpA	axial spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASDAI50	Bath Ankylosing Spondylitis Disease Activity Index 50% response criteria
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BMI	body mass index
BMP	bone morphogenic protein
BP	bodily pain
CI	confidence interval
CK	creatinine kinase
CR	compliance ratio
CRP	C-reactive protein
CSR	clinical study report
CV	coefficient of variation
CZP	certolizumab pegol
DEM	Data Evaluation Meeting
DKK1	dickkopf-related protein 1
DRL	Drug Reference List
DMARD	disease-modifying antirheumatic drug
EAIR	exposure-adjusted incidence rate

---

EAER	exposure-adjusted event rate
eCRF	electronic case report form
EQ-5D	EuroQoL Health Status Questionnaire (5 dimensions)
ER	emergency room
ES	Enrolled Set
FAS	Full Analysis Set
GH	general health
HCQ	hydroxychloroquine
HLA-B27	human leukocyte antigen B27
HLT	High Level Term
HRQoL	health-related quality of life
ICH	International Conference on Harmonisation
IXRS	interactive response system
IBD	inflammatory bowel disease
LOCF	last observation carried forward
LLN	lower limit of normal
LLOQ	lower limit of quantification
LLT	Low Level Term
MASES	Maastricht Ankylosing Spondylitis Enthesitis Score
MCS	Mental Component Summary
MedDRA <sup>®</sup>	Medical Dictionary of Regulatory Activities <sup>®</sup>
MH	mental health
MMP	matrix metalloproteinase
mNY	Modified New York (criteria)

---

MOS	Medical Outcomes Study
MRI	magnetic resonance imaging
MTX	methotrexate
nADA	neutralizing ADA <sub>b</sub>
nr-axSpA	nonradiographic axSpA
NRS	Numeric Rating Scale
NRI	non-responder imputation
NSAID	nonsteroidal anti-inflammatory drug
OC	observed case
OL	open-label
OL-CZP	open-label CZP
OL-OT	open-label other treatment
PBO	placebo
PCS	Physical Component Summary
PF	physical function
PGADA	Patient's Global Assessment of Disease Activity
PhGADA	Physician's Global Assessment of Disease Activity
PPS	Per Protocol Set
PT	Preferred Term
Q	question
Q1	1 <sup>st</sup> Quartile
Q3	3 <sup>rd</sup> Quartile
Q2W	every 2 weeks (every other week)
QoL	quality of life

---

RBC	red blood cell
RCTC	Rheumatology Common Toxicity Criteria
RE	role emotional
RP	role physical
RS	Randomized Set
SAARD	slow-acting anti-rheumatic drug
SAP	statistical analysis plan
sc	subcutaneous(ly)
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SF	social functioning
SFE	safety follow-up extension
SFE-WD	safety follow-up extension-early withdrawal visit
SFE-FAS	safety follow-up extension full analysis set
SF-36	Short-Form 36-Item Health Survey
SI	sacroiliac
SJC	swollen joint count
SM	spinal mobility
SMQ	Standardized MedDRA Query
SOC	System Organ Class
SOP	standard operating procedure
SpA	spondyloarthritis
SPARCC	Spondyloarthritis Research Consortium of Canada

---

sqrt	square root
SS	Safety Set
SSZ	sulfasalazine
STIR	short-tau-inversion recovery
TB	tuberculosis
TEAE	treatment-emergent adverse event
TJC	tender joint count
TNF	tumor necrosis factor
USA	United States of America
ULN	upper limit of normal
VAS	visual analog scale
VT	vitality
VU	vertebral units
WBC	white blood cell
WHO	World Health Organization
WPS	Work Productivity Survey

## Change #2

[Section 2.3.1](#) was change to add “but a  $\pm 4$  week window will be used in the analysis” to the third to last paragraph. The paragraph was changed from:

Starting at Week 64, all subjects are to visit the site approximately every 12 weeks ( $\pm 2$  weeks) for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156/Safety Follow-up Extension-Early Withdrawal Visit (SFE-WD) ( $\pm 1$  week).

To:

Starting at Week 64, all subjects are to visit the site approximately every 12 weeks ( $\pm 2$  weeks) for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156/Safety Follow-up Extension-Early Withdrawal Visit (SFE-WD) ( $\pm 1$  week but a  $\pm 4$  week window will be used in the analysis).

## Change #3

[Section 3.2.1.5](#) was changed from:

### 3.2.1.5 SFE period

The SFE period starts when the subject takes their first dose of OL-CZP that is after week 50 and after date of informed consent for the SFE. Visit based efficacy or safety data that is recorded on the day of the first SFE period dose will be assigned to the previous period unless available time information clearly indicates that it was recorded after the first administration of SFE period treatment.

To:

### 3.2.1.5 Pre-SFE period

The pre-SFE period starts after last dose of the double-blind period (i.e. also accounts for the last OL dose if the subject escaped) + 70 days (note that the last double-blind dose is at Week 50, thus each AE which starts up to 70 days after the last double-blind dose is still assigned to the double-blind phase) and ends before SFE period starts. The first SFE dose is used to determine if AEs end before the SFE period starts since this is the time when the AEs become treatment-emergent.

#### Change #4

The original [Section 3.2.1.5](#) was changed to [section 3.2.1.6](#) and the text was changed from: The SFE period starts when the subject takes their first dose of OL-CZP that is after week 50 and after date of informed consent for the SFE. Visit based efficacy or safety data that is recorded on the day of the first SFE period dose will be assigned to the previous period unless available time information clearly indicates that it was recorded after the first administration of SFE period treatment.

The SFE period ends when the subject discontinues from or completes the study following SFE enrollment (Week 156/SFE-WD).

To: The SFE period starts when the subject takes their first dose of OL-CZP that is after week 50 and after date of informed consent for the SFE. Visit based efficacy or safety data that is recorded on the day of the first SFE period dose will be assigned to the previous period unless available time information clearly indicates that it was recorded after the first administration of SFE period treatment. There are some subjects who took their first dose of OL-CZP that is after week 50 before the informed consent for the SFE. In such cases the SFE period will be programmed to start at the informed consent date.

The SFE period ends when the subject discontinues from or completes the study following SFE enrollment (Week 156/SFE-WD).

#### Change #5

[Section 3.3](#) was changed to add a final paragraph and clarify that the same baseline definition is used in the SFE. It was changed from:

Unless otherwise specified, the last valid measurement before study medication administration in the double-blind period will be used as the Baseline value. The same Baseline definition will be used for the open-label period. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables (eg, demography) assessments are scheduled for the Screening visit only and not for the Baseline visit. In this case, the Screening value will be utilized as Baseline value. 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.

Baseline images, either spinal x-ray or MRI assessments, are required to evaluate study eligibility. Therefore, all Baseline images will be available prior to randomization into the study. SI joint x-rays should not be older than 12 months prior to Baseline and should be verified by central reading during the Screening period.

To:

Unless otherwise specified, the last valid measurement before study medication administration in the double-blind period will be used as the Baseline value. The same Baseline definition will be used for the open-label period and SFE period. For almost all variables, this will be the assessment made at the Baseline visit. However, for some variables (eg, demography) assessments are scheduled for the Screening visit only and not for the Baseline visit. In this case, the Screening value will be utilized as Baseline value. 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.

Baseline images, either spinal x-ray or MRI assessments, are required to evaluate study eligibility. Therefore, all Baseline images will be available prior to randomization into the study. SI joint x-rays should not be older than 12 months prior to Baseline and should be verified by central reading during the Screening period.

There will be an additional SFE Baseline presented in some tables. It is defined as the last value before the patient received SFE treatment (the first dose on or after SFE informed consent, for most patients this will be the Week 52 value).

#### **Change #6**

The first sentence was removed from the last paragraph in Section 3.2.2 which was changed from:

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

To:

For AEs, relative days for start and stop dates will be calculated as the number of days since the first injection of the relevant double-blind or OL-CZP medication. For non-treatment emergent AEs, relative day of onset will be negative if the event started and stopped before the first dose.

#### **Change #7**

[Section 3.9](#) was updated to add the second paragraph. It was changed from:

The protocol stated that ADAb concentrations above 2.4 unit/ml would be defined as ADAb positive. However, during the trial it was noted that the background sample noise could only be determined during the sample analyses and hence the cut point cannot be pre-specified. Therefore the immunogenicity sections of the SAP were updated.

To:

The protocol stated that ADAb concentrations above 2.4 unit/ml would be defined as ADAb positive. However, during the trial it was noted that the background sample noise could only be determined during the sample analyses and hence the cut point cannot be pre-specified. Therefore the immunogenicity sections of the SAP were updated.

The protocol defines a plus/minus one week window for the Week 156 time point. However a plus/minus 4 week window will be used in the analysis due to Corona Virus Disease 2019 pandemic.

### Change #8

The section and third paragraph in [Section 4.2.3](#) was changed from:

For analysis of the Week 156/SFE-WD visit data, binary efficacy variables will be treated in the same way as the primary analysis of the primary efficacy variable (Composite Endpoint). That is, a subject will be considered a responder only if the response is achieved and the last administration of OL-CZP study medication is on or after Week 154/Day 1079 (post randomization). Subjects with missing data at Week 156/SFE-WD; or who have withdrawn treatment prior to Day 1079; or those without response at Week 156/SFE-WD, are counted as non-responders. Hence all subjects in the SFE-FAS analysis set will be included in this analysis.

Two observed case sensitivity analyses of binary efficacy variables at Week 156/SFE-WD will be performed. The Observed Case analysis at Week 156, will consider a subject a responder only if the response is achieved at the Week 156 visit and the subject completed the study. Non-response at Week 156 visit will be counted as non-responders. Early withdrawals or subjects with missing data will be considered missing in this analysis. The Observed Case analysis at Week 156/SFE-WD will consider a subject as a responder if response is achieved at Week 156/SFE-WD visit such that both subjects with a Week 156 visit (subjects who complete) and subjects with early withdrawal visits and efficacy results meeting the criteria of response, will be counted as a responder. Subjects with missing data will be considered missing in this analysis.

To:

For analysis of the Week 156/SFE-WD visit data, binary efficacy variables will be analyzed using the non-responder imputation (NRI) approach; that is, where the denominator is all subjects in the SFE-FAS for the given treatment group. A subject will be considered a responder only if the response criterion of the given variable is achieved. All other subjects will be counted as not responding e.g. subjects who have not achieved the given outcome and all subjects with missing data are set to non-response.

An observed case analysis of binary efficacy variables at Week 156/SFE-WD will be performed. The Observed Case analysis at Week 156, will consider a subject a responder only if the response is achieved. Non-response at Week 156 visit will be counted as non-responders. Early withdrawals or subjects with missing data will be considered missing in this analysis. The sensitivity analysis of the Observed Case analysis at Week 156/SFE-WD will consider a subject

as a responder if response is achieved at Week 156/SFE-WD visit such that both subjects with a Week 156 visit (subjects who complete) and subjects with early withdrawal visits and efficacy results meeting the criteria of response, will be counted as a responder. Subjects with missing data will be considered missing in this analysis.

### Change #9

Additional imputation rules for completely missing AE dates were added in Section 4.2.4 which was changed from:

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. If time of study treatment and time of event or concomitant medication is available and non partial then it will also be used however if missing or partial, only the dates will be compared. 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 double-blind 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 administration of study medication during the double-blind period. 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 double-blind 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 administration of double-blind 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.

- 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 administration 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 administration of study medication. Otherwise the end day will be set to December 31st of the given year.

To

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. If time of study treatment and time of event or concomitant medication is available and non partial then it will also be used however if missing or partial, only the dates will be compared. 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.

- Missing start day, but month and year present:  
If the start of double-blind 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 administration of study medication during the double-blind period. 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 double-blind 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 administration of double-blind study medication. Otherwise the start day and month will be set to January 1st.
- Missing start day, month, and year:  
The start date will be replaced by the date of first study drug administration of double-blind study medication or by the first day of January of the year of the end date, if the first administration date of double-blind study medication 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 administration 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 administration of study medication. Otherwise the end day will be set to December 31st of the given year.
- If the stop date is completely unknown, the stop date will not be imputed.

#### Change#10

The second paragraph in [Section 4.3.2](#) was updated from:

For analysis of the Week 156/SFE-WD visit data, binary efficacy variables will be treated in the same way as the primary analysis of the primary efficacy variable (Composite Endpoint). That is, a subject will be considered a responder only if the response is achieved and the last administration of OL CZP study medication is on or after Week 154/Day 1079 (post-randomization). Subjects with missing data at Week 156/SFE-WD; or who have withdrawn treatment prior to Day 1079; or those without response at Week 156/SFE-WD, are counted as non-responders. Hence all subjects in the SFE-FAS analysis set will be included in this analysis.

To:

For analysis of the Week 156/SFE-WD visit data, binary efficacy variables will be analyzed using the NRI approach; that is, where the denominator is all subjects in the SFE-FAS for the given treatment group. A subject will be considered a responder only if the response criterion of the given variable is achieved. All subjects with missing data are set to non-response i.e. all other subjects will be counted as not responding e.g. subjects who have not achieved the given outcome and all subjects with missing data are set to non-response.

### Change #11

[Section 6.4](#) was updated to add text to the penultimate paragraph. It was changed from:

Concomitant DMARDs and NSAIDs will be summarized by WHO-DRL ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and Preferred Term).

To:

Concomitant DMARDs and NSAIDs will be summarized by WHO-DRL ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and Preferred Term) for the study as a whole and separately for those medications taken during the SFE for the SFE Safety Set.

### Change #12

The second set of bullets in [Section 8.2.1.4](#) was updated from:

For post-Baseline visits, a time window of 2 weeks before or after the scheduled visit will be used for mapping MRI data to the given visit.

To:

- With the exception of Week 156, for post-Baseline visits, a time window of 2 weeks before or after the scheduled visit will be used for mapping MRI data to the given visit.
- For Week 156, a time window of 4 weeks before or after the scheduled visit will be used for MRI data to the given visit.

### Change #13

The third paragraph in [Section 8.3](#) was changed from:

Categorical other efficacy variables will be analyzed using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+). If the logistic regression model is unable to converge, then the MRI/CRP classification variable may be dropped from the model to facilitate convergence. Tables for the binary variables will display the responder rates for the various time points. Analyses will be presented in 2 ways (as described in [Section 4.2.3](#)). First, they will be presented where the denominator is all subjects in the FAS for the given treatment group. This is essentially a non-response imputation (NRI) approach as subjects who have not achieved the given outcome (whether observed or not) or subjects who discontinue double-blind study treatment are considered as not having responded. Second, the responder rates will be presented for the FAS using Observed Case data. Subjects still on the original double-blind study treatment and with response will be analyzed, with the denominator based on the number of observed values at the given time point.

To:

Categorical other efficacy variables will be analyzed using a logistic regression model with factors of treatment group, region, and MRI/CRP classification (MRI+/CRP+; MRI+/CRP-; MRI-/CRP+). If the logistic regression model is unable to converge, then the MRI/CRP classification variable may be dropped from the model to facilitate convergence. Tables for the binary variables will display the responder rates for the various time points. Analyses will be presented in 2 ways (as described in [Section 4.2.3](#)). First, they will be presented where the

denominator is all subjects in the FAS for the given treatment group. This is essentially a NRI approach as subjects who have not achieved the given outcome (whether observed or not) or subjects who discontinue double-blind study treatment are considered as not having responded. Second, the responder rates will be presented for the FAS using Observed Case data. Subjects still on the original double-blind study treatment and with response will be analyzed, with the denominator based on the number of observed values at the given time point.

#### **Change #14**

The following last paragraph was removed from [Section 10.3](#):

A summary table of all TEAEs (by SOC, HLT and PT) by TE ADAb Status. For this summary, subjects will be categorized using the treatment-emergent ADAb classification, and will be presented for TEAEs occurring prior to becoming treatment emergent ADAb positive, TEAEs occurring after becoming TE ADAb positive, and TEAEs for subjects who remained TE ADAb negative. This summary may be produced using a titer threshold (to be determined from the data) instead of or in addition to TE ADAb Status.

#### **Change #15**

The following sentence was added to the end of [Section 11.1](#):

The following doses will be presented in a listing:

- very first dose in the study, the very last dose in the study and the difference between them
- last dose in the DB period (i.e. last DB dose or last OL dose), the last SFE dose and the difference between them.

#### **Change #16**

The first paragraph in [Section 11.2](#) was changed from:

AE start dates will be examined to determine if they are treatment-emergent and attributable to the double-blind period, open-label period or SFE period. AEs occurring either before double-blind study drug administration (pre-treatment) or outside of the windows described below will be defined as non-treatment emergent.

To:

AE start dates will be examined to determine if they are treatment-emergent and attributable to the double-blind period, open-label period, pre-SFE period or SFE period. AEs occurring either before double-blind study drug administration (pre-treatment) or outside of the windows described below will be defined as non-treatment emergent.

**Change #17**

Table 11-1 was changed from:

<b>Table 14–1: Assignment of Adverse Events to Study Periods</b>		
<b>Type of Subject</b>	<b>Assignment Rule Based on AE Start Date</b>	<b>Assignment</b>
Subjects who do not enter the open-label period and do not enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days.	Double-blind TEAE
Subjects who do not enter the open-label period but do enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and prior to loading dose/first dose administration in SFE period	Double-blind TEAE
Subjects who enter the open-label period	1. Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and before the start date and time of OL-CZP or OL-OT.  <b>Note:</b> If a subject continues in the study with OL-OT visits but without taking any relevant other medication, there will be no start date of the other treatment period and all AEs with a start date within 70 days after last administration will be assigned to the DB period.	Double-blind TEAE
	2. Start date is on or after the start date of OL-CZP and on or before the last study drug administration in the OL-CZP period + 70 days and <ul style="list-style-type: none"> <li>• before the start date of OL-OT (for subjects moving into OL-OT), or</li> <li>• before the first dose administration in SFE period (for subjects moving into SFE period)</li> </ul>	Open-label CZP TEAE
	3. Start date is on or after the start date of OL-OT.  <b>Note:</b> Other treatment adverse events will not be summarized, but only presented in listings	Open-label other treatment TEAE
Subjects who enter the open-label SFE period	Start date is on or after the loading dose/first dose administration in SFE period and on or before the last study drug administration in the SFE period + 70 days.	SFE TEAE

To:

<b>Table 14–1: Assignment of Adverse Events to Study Periods</b>		
<b>Type of Subject</b>	<b>Assignment Rule Based on AE Start Date</b>	<b>Assignment</b>
Subjects who do not enter the open-label period and do not enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days.	Double-blind TEAE
Subjects who do not enter the open-label period but do enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and prior to loading dose/first dose administration in SFE period	Double-blind TEAE
Subjects who enter the open-label period	1. Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and before the start date and time of OL-CZP or OL-OT.  <b>Note:</b> If a subject continues in the study with OL-OT visits but without taking any relevant other medication, there will be no start date of the other treatment period and all AEs with a start date within 70 days after last administration will be assigned to the DB period.	Double-blind TEAE
	2. Start date is on or after the start date of OL-CZP and on or before the last study drug administration in the OL-CZP period + 70 days and <ul style="list-style-type: none"> <li>• before the start date of OL-OT (for subjects moving into OL-OT), or</li> <li>• before the first dose administration in SFE period (for subjects moving into SFE period)</li> </ul>	Open-label CZP TEAE
	3. Start date is on or after the start date of OL-OT.  <b>Note:</b> Other treatment adverse events will not be summarized, but only presented in listings	Open-label other treatment AE
Subjects who enter the pre-SFE period	Start date is after the last double-blind dose + 70 days and is before the loading dose/first dose administration in SFE period.	Pre-SFE AE

<b>Table 14–1: Assignment of Adverse Events to Study Periods</b>		
<b>Type of Subject</b>	<b>Assignment Rule Based on AE Start Date</b>	<b>Assignment</b>
Subjects who enter the open-label SFE period	Start date is on or after the loading dose/first dose administration in SFE period and on or before the last study drug administration in the SFE period + 70 days.	SFE TEAE

**Change #18**

The following bulleted list in Section 11.2 was changed from:

The following AE summaries will be presented:

- Overview of TEAE
- TEAE
- Serious TEAE
- Non-serious TEAE
- TEAE with fatal outcome
- TEAE leading to permanent withdrawal of study medication
- Severe TEAE (included in TEAE table by intensity)
- Drug-related TEAE (included in TEAE table by relationship)
- Serious TEAE by relationship
- Non-serious TEAE by relationship
- Fatal TEAE by relationship
- TEAE of special interest (see below)
- TEAE sorted by incidence of PTs in the CZP 200mg group (SOC and higher level term [HLT] will not be utilized)
- Non-serious TEAE with an incidence of more than 5% in 1 of the 2 randomized groups (HLT will not be utilized)
- Non-serious TEAE with an incidence of more than 5% in 1 of the 2 randomized groups by relationship (HLT will not be utilized)
- TEAE subject numbers
- TEAEs with a CZP incidence  $\geq 1\%$  and CZP incidence  $>$  PBO incidence

To:

The following AE summaries will be presented:

- Overview of TEAE
- Overview of Pre-SFE AEs

- TEAE
- Serious TEAE
- Non-serious TEAE
- TEAE with fatal outcome
- TEAE leading to permanent withdrawal of study medication
- Severe TEAE (included in TEAE table by intensity)
- Drug-related TEAE (included in TEAE table by relationship)
- Serious TEAE by relationship
- Non-serious TEAE by relationship
- Fatal TEAE by relationship
- TEAE of special interest (see below)
- TEAE sorted by incidence of PTs in the CZP 200mg group (SOC and higher level term [HLT] will not be utilized)
- Non-serious TEAE with an incidence of more than 5% in 1 of the 2 randomized groups (HLT will not be utilized)
- Non-serious TEAE with an incidence of more than 5% in 1 of the 2 randomized groups by relationship (HLT will not be utilized)
- TEAE subject numbers
- TEAEs with a CZP incidence  $\geq 1\%$  and CZP incidence  $>$  PBO incidence

#### **Change #19**

The following paragraph in [Section 11.2](#) was changed from:

The TEAE incidence rates will also be adjusted for exposure (as calculated using methods 3 and 4 in [Section 11.1](#)) and reported by 100 subject years of exposure. 100 subject years is defined as the sum of the exposure / number of subjects \* 100. Two approaches will be applied to adjust for exposure. One will only use the first occurrence of an AE with corresponding exposure (EAIR) and the second approach will use all AEs and the entire exposure (EAER).

To:

The TEAE incidence rates will also be adjusted for exposure (as calculated using methods 3 and 4 in [Section 11.1](#)) and reported by 100 subject years of exposure. Two approaches will be applied to adjust for exposure. One will only use the first occurrence of an AE with corresponding exposure (EAIR) and the second approach will use all AEs and the entire exposure (EAER).

## 14.6 AMENDMENT 6

### Rationale for the amendment

The analysis plan was mainly updated to adjust the windowing rule for Week 156. All modifications are detailed below.

### Change #1

The list of abbreviations was updated from:

ADAb	anti-drug antibody
ADaM	analysis data model
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AP	alkaline phosphatase
ASAS	Assessment of SpondyloArthritis international Society
ASAS20, 40	Assessment of SpondyloArthritis international Society 20%, 40% response criteria
ASAS5/6	Assessment of SpondyloArthritis international Society 20% improvement in 5 of 6 domains
AS	ankylosing spondylitis
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASDAS-CII	Ankylosing Spondylitis Disease Activity Score – Clinically Important Improvement
ASDAS-HD	Ankylosing Spondylitis Disease Activity Score – High Disease activity
ASDAS-ID	Ankylosing Spondylitis Disease Activity Score – Inactive Disease
ASDAS-LD	Ankylosing Spondylitis Disease Activity Score – Low Disease
ASDAS-MD	Ankylosing Spondylitis Disease Activity Score – Moderate Disease

---

ASDAS-MI	Ankylosing Spondylitis Disease Activity Score – Major Improvement
ASDAS-vHD	Ankylosing Spondylitis Disease Activity Score – very High Disease activity
ASQoL	Ankylosing Spondylitis Quality of Life
ASspiMRI-a	Ankylosing Spondylitis spine MRI acuity
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AU	anterior uveitis
AxSpA	axial spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASDAI50	Bath Ankylosing Spondylitis Disease Activity Index 50% response criteria
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BMI	body mass index
BMP	bone morphogenic protein
BP	bodily pain
CI	confidence interval
CK	creatinine kinase
CR	compliance ratio
CRP	C-reactive protein
CSR	clinical study report
CV	coefficient of variation
CZP	certolizumab pegol
DEM	Data Evaluation Meeting

---

DKK1	dickkopf-related protein 1
DRL	Drug Reference List
DMARD	disease-modifying antirheumatic drug
EAIR	exposure-adjusted incidence rate
EAER	exposure-adjusted event rate
eCRF	electronic case report form
EQ-5D	EuroQoL Health Status Questionnaire (5 dimensions)
ER	emergency room
ES	Enrolled Set
FAS	Full Analysis Set
GH	general health
HCQ	hydroxychloroquine
HLA-B27	human leukocyte antigen B27
HLT	High Level Term
HRQoL	health-related quality of life
ICH	International Conference on Harmonisation
IXRS	interactive response system
IBD	inflammatory bowel disease
LOCF	last observation carried forward
LLN	lower limit of normal
LLOQ	lower limit of quantification
LLT	Low Level Term
MASES	Maastricht Ankylosing Spondylitis Enthesitis Score
MCS	Mental Component Summary

---

MedDRA®	Medical Dictionary of Regulatory Activities®
MH	mental health
MMP	matrix metalloproteinase
mNY	Modified New York (criteria)
MOS	Medical Outcomes Study
MRI	magnetic resonance imaging
MTX	methotrexate
nADA	neutralizing ADA <sub>b</sub>
nr-axSpA	nonradiographic axSpA
NRS	Numeric Rating Scale
NRI	non-responder imputation
NSAID	nonsteroidal anti-inflammatory drug
OC	observed case
OL	open-label
OL-CZP	open-label CZP
OL-OT	open-label other treatment
PBO	placebo
PCS	Physical Component Summary
PF	physical function
PGADA	Patient's Global Assessment of Disease Activity
PhGADA	Physician's Global Assessment of Disease Activity
PPS	Per Protocol Set
PT	Preferred Term
Q	question

---

Q1	1 <sup>st</sup> Quartile
Q3	3 <sup>rd</sup> Quartile
Q2W	every 2 weeks (every other week)
QoL	quality of life
RBC	red blood cell
RCTC	Rheumatology Common Toxicity Criteria
RE	role emotional
RP	role physical
RS	Randomized Set
SAARD	slow-acting anti-rheumatic drug
SAP	statistical analysis plan
sc	subcutaneous(ly)
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SF	social functioning
SFE	safety follow-up extension
SFE-WD	safety follow-up extension-early withdrawal visit
SFE-FAS	safety follow-up extension full analysis set
SF-36	Short-Form 36-Item Health Survey
SI	sacroiliac
SJC	swollen joint count
SM	spinal mobility
SMQ	Standardized MedDRA Query

---

SOC	System Organ Class
SOP	standard operating procedure
SpA	spondyloarthritis
SPARCC	Spondyloarthritis Research Consortium of Canada
sqrt	square root
SS	Safety Set
SSZ	sulfasalazine
STIR	short-tau-inversion recovery
TB	tuberculosis
TEAE	treatment-emergent adverse event
TJC	tender joint count
TNF	tumor necrosis factor
USA	United States of America
ULN	upper limit of normal
VAS	visual analog scale
VT	vitality
VU	vertebral units
WBC	white blood cell
WHO	World Health Organization
WPS	Work Productivity Survey

To:

ADAb	anti-drug antibody
ADaM	analysis data model
AE	adverse event

ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AP	alkaline phosphatase
ASAS	Assessment of SpondyloArthritis international Society
ASAS20, 40	Assessment of SpondyloArthritis international Society 20%, 40% response criteria
ASAS5/6	Assessment of SpondyloArthritis international Society 20% improvement in 5 of 6 domains
AS	ankylosing spondylitis
ASDAS	Ankylosing Spondylitis Disease Activity Score
ASDAS-CII	Ankylosing Spondylitis Disease Activity Score – Clinically Important Improvement
ASDAS-HD	Ankylosing Spondylitis Disease Activity Score – High Disease activity
ASDAS-ID	Ankylosing Spondylitis Disease Activity Score – Inactive Disease
ASDAS-LD	Ankylosing Spondylitis Disease Activity Score – Low Disease
ASDAS-MD	Ankylosing Spondylitis Disease Activity Score – Moderate Disease
ASDAS-MI	Ankylosing Spondylitis Disease Activity Score – Major Improvement
ASDAS-vHD	Ankylosing Spondylitis Disease Activity Score – very High Disease activity
ASQoL	Ankylosing Spondylitis Quality of Life
ASspiMRI-a	Ankylosing Spondylitis spine MRI acuity
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AU	anterior uveitis

---

AxSpA	axial spondyloarthritis
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASDAI50	Bath Ankylosing Spondylitis Disease Activity Index 50% response criteria
BASFI	Bath Ankylosing Spondylitis Functional Index
BASMI	Bath Ankylosing Spondylitis Metrology Index
BMI	body mass index
BMP	bone morphogenic protein
BP	bodily pain
CI	confidence interval
CK	creatinine kinase
COVID-19	Corona Virus Disease 2019
CR	compliance ratio
CRP	C-reactive protein
CSR	clinical study report
CV	coefficient of variation
CZP	certolizumab pegol
DEM	Data Evaluation Meeting
DKK1	dickkopf-related protein 1
DRL	Drug Reference List
DMARD	disease-modifying antirheumatic drug
EAIR	exposure-adjusted incidence rate
EAER	exposure-adjusted event rate
eCRF	electronic case report form
EQ-5D	EuroQoL Health Status Questionnaire (5 dimensions)

---

ER	emergency room
ES	Enrolled Set
FAS	Full Analysis Set
GH	general health
HCQ	hydroxychloroquine
HLA-B27	human leukocyte antigen B27
HLT	High Level Term
HRQoL	health-related quality of life
ICH	International Conference on Harmonisation
IXRS	interactive response system
IBD	inflammatory bowel disease
LOCF	last observation carried forward
LLN	lower limit of normal
LLOQ	lower limit of quantification
LLT	Low Level Term
MASES	Maastricht Ankylosing Spondylitis Enthesitis Score
MCS	Mental Component Summary
MedDRA <sup>®</sup>	Medical Dictionary of Regulatory Activities <sup>®</sup>
MH	mental health
MMP	matrix metalloproteinase
mNY	Modified New York (criteria)
MOS	Medical Outcomes Study
MRI	magnetic resonance imaging
MTX	methotrexate

---

nADA	neutralizing ADA <sub>b</sub>
nr-axSpA	nonradiographic axSpA
NRS	Numeric Rating Scale
NRI	non-responder imputation
NSAID	nonsteroidal anti-inflammatory drug
OC	observed case
OL	open-label
OL-CZP	open-label CZP
OL-OT	open-label other treatment
PBO	placebo
PCS	Physical Component Summary
PF	physical function
PGADA	Patient's Global Assessment of Disease Activity
PhGADA	Physician's Global Assessment of Disease Activity
PPS	Per Protocol Set
PT	Preferred Term
Q	question
Q1	1 <sup>st</sup> Quartile
Q3	3 <sup>rd</sup> Quartile
Q2W	every 2 weeks (every other week)
QoL	quality of life
RBC	red blood cell
RCTC	Rheumatology Common Toxicity Criteria
RE	role emotional

---

RP	role physical
RS	Randomized Set
SAARD	slow-acting anti-rheumatic drug
SAP	statistical analysis plan
sc	subcutaneous(ly)
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SF	social functioning
SFE	safety follow-up extension
SFE-WD	safety follow-up extension-early withdrawal visit
SFE-FAS	safety follow-up extension full analysis set
SF-36	Short-Form 36-Item Health Survey
SI	sacroiliac
SJC	swollen joint count
SM	spinal mobility
SMQ	Standardized MedDRA Query
SOC	System Organ Class
SOP	standard operating procedure
SpA	spondyloarthritis
SPARCC	Spondyloarthritis Research Consortium of Canada
sqrt	square root
SS	Safety Set
SSZ	sulfasalazine

---

STIR	short-tau-inversion recovery
TB	tuberculosis
TEAE	treatment-emergent adverse event
TJC	tender joint count
TNF	tumor necrosis factor
USA	United States of America
ULN	upper limit of normal
VAS	visual analog scale
VT	vitality
VU	vertebral units
WBC	white blood cell
WHO	World Health Organization
WPS	Work Productivity Survey

### Change #2

The following paragraph in [Section 2.3.1](#) was changed from:

Starting at Week 64, all subjects are to visit the site approximately every 12 weeks ( $\pm 2$  weeks) for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156/Safety Follow-up Extension-Early Withdrawal Visit (SFE-WD) ( $\pm 1$  week but a  $\pm 4$  week will be used in the analysis.

To:

Starting at Week 64, all subjects are to visit the site approximately every 12 weeks ( $\pm 2$  weeks) for assessments. The last dosing visits will be at Week 154. The final study assessments are performed at Week 156/Safety Follow-up Extension-Early Withdrawal Visit (SFE-WD) ( $\pm 1$  week). However a  $\pm 4$  week window will be used in the analysis for the Week 156 timepoint). The expected day for Week 156 will be calculated assuming that the SFE period start was on Week 52, i.e. Week 156 will be expected to be  $(156-52) \times 7$  days after Week 52. For patients where the first dose after the Week 50 visit is before the SFE informed consent the width of the lower boundary of the Week 156 window will be increased further by the absolute number of days between the first dose after the Week 50 visit and the SFE informed consent date.

### Change #3

The second paragraph in [Section 3.9](#) was changed from:

The protocol defines a plus/minus one week window for the Week 156 time point. However a plus/minus 4 week window will be used in the analysis due to the Corona Virus Disease 2019 pandemic.

To:

The protocol defines a plus/minus one week window for the Week 156 time point. However at least a plus/minus 4 week window will be used in the analysis due to the Corona Virus Disease 2019 (COVID-19) pandemic, with an even more generous window for patients who had their first dose post-Week 50 before the SFE informed consent date.

### Change #4

The first and second bullet in [Section 4.2.4](#) was changed from:

- Missing start day, but month and year present:

If the start of double-blind 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 administration of study medication during the double-blind period. 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 double-blind 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 administration of double-blind study medication. Otherwise the start day and month will be set to January 1st.

To:

- Missing start day, but month and year present:

If the start of double-blind 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 administration of study medication during the double-blind period. Otherwise if informed consent was signed 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 informed consent.

Otherwise if the following conditions are met:

- the year of the occurrence of the AE/concomitant medication is after the year of the end of the double-blind period or
- the year of the occurrence of the AE/concomitant medication is the same as the year of the end of the double-blind period and the month of the occurrence of the AE/concomitant medication is the after the month of the end of the double-blind period

then the following imputations will be done:

If the year of the occurrence of the AE/concomitant medication is the same as the start of the open-label period and the month of the AE/concomitant medication is the same as the month of the start of the open-label period then the start day of the event/concomitant medication will be assigned to the first day of the open-label period.

Otherwise if the year of the occurrence of the AE/concomitant medication is the same as the start of the SFE period and the month of the AE/concomitant medication is the same as the month of the start of the SFE period then the start day of the event/concomitant medication will be assigned to the first day of the SFE period.

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 double-blind 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 administration of double-blind study medication. Otherwise if informed consent was signed in the same year as the occurrence of the AE/concomitant medication, the start day and month will be assigned to the day and month of informed consent.

Otherwise if the year of the occurrence of the AE/concomitant medication is after the year of the end of the double-blind period and then the following imputations will be done:

If the year of the occurrence of the AE/concomitant medication is the same as the start of the open-label period then the start day and month of the event/concomitant medication will be assigned to the first day of the open-label period.

Otherwise if the year of the occurrence of the AE/concomitant medication is the same as the start of the SFE period then the start day of the event/concomitant medication will be assigned to the first day of the SFE period.

Otherwise the start day and month will be set to January 1st.

#### Change #5

The second set of bullets in [Section 8.2.1.4](#) was updated from:

- All Baseline values should be collected prior to randomization as MRI is required to be randomized into the study.
- With the exception of Week 156, for post-Baseline visits, a time window of 2 weeks before or after the scheduled visit will be used for mapping MRI data to the given visit.
- For Week 156, a time window of 4 weeks before or after the scheduled visit will be used for MRI data to the given visit.

To:

- All Baseline values should be collected prior to randomization as MRI is required to be randomized into the study.
- With the exception of Week 156, for post-Baseline visits, a time window of 2 weeks before or after the scheduled visit will be used for mapping MRI data to the given visit.

- For Week 156, a time window of 4 weeks before or after the scheduled visit will be used for MRI data to the given visit. The expected day for Week 156 will be calculated assuming that the SFE period start was on Week 52. I.e. Week 156 will be expected to be  $(156-52) \times 7$  days after Week 52. For patients where the first dose after the Week 50 is before the SFE informed consent the width of the lower boundary of the Week 156 window will be increased further by the absolute number of days between the first dose after the Week 50 visit and the SFE informed consent date.

**Change #6**

The following table was updated for consistency.

From:

<b>Table 14–1: Assignment of Adverse Events to Study Periods</b>		
<b>Type of Subject</b>	<b>Assignment Rule Based on AE Start Date</b>	<b>Assignment</b>
Subjects who do not enter the open-label period and do not enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days.	Double-blind TEAE
Subjects who do not enter the open-label period but do enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and prior to loading dose/first dose administration in SFE period	Double-blind TEAE

**Table 14–1: Assignment of Adverse Events to Study Periods**

Type of Subject	Assignment Rule Based on AE Start Date	Assignment
Subjects who enter the open-label period	<p>1. Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and before the start date and time of OL-CZP or OL-OT.</p> <p><b>Note:</b> If a subject continues in the study with OL-OT visits but without taking any relevant other medication, there will be no start date of the other treatment period and all AEs with a start date within 70 days after last administration will be assigned to the DB period.</p>	Double-blind TEAE
	<p>2. Start date is on or after the start date of OL-CZP and on or before the last study drug administration in the OL-CZP period + 70 days and</p> <ul style="list-style-type: none"> <li>• before the start date of OL-OT (for subjects moving into OL-OT), or</li> <li>• before the first dose administration in SFE period (for subjects moving into SFE period)</li> </ul>	Open-label CZP TEAE
	<p>3. Start date is on or after the start date of OL-OT.</p> <p><b>Note:</b> Other treatment adverse events will not be summarized, but only presented in listings</p>	Open-label other treatment TEAE
Subjects who enter the pre-SFE period	Start date is after the last double-blind dose + 70 days and is before the loading dose/first dose administration in SFE period.	Pre-SFE AE
Subjects who enter the open-label SFE period	Start date is on or after the loading dose/first dose administration in SFE period and on or before the last study drug administration in the SFE period + 70 days.	SFE TEAE

To:

**Table 14–1: Assignment of Adverse Events to Study Periods**

Type of Subject	Assignment Rule Based on AE Start Date	Assignment
Subjects who do not enter the open-label period and do not enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days.	Double-blind TEAE
Subjects who do not enter the open-label period but do enter SFE period	Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and prior to the SFE period start	Double-blind TEAE
Subjects who enter the open-label period	1. Start date is on or after first study drug administration during the double-blind period and on or before the last study drug administration in the double-blind period + 70 days and before the start date and time of OL-CZP or OL-OT.  <b>Note:</b> If a subject continues in the study with OL-OT visits but without taking any relevant other medication, there will be no start date of the other treatment period and all AEs with a start date within 70 days after last administration will be assigned to the DB period.	Double-blind TEAE
	2. Start date is on or after the start date of OL-CZP and on or before the last study drug administration in the OL-CZP period + 70 days and <ul style="list-style-type: none"> <li>• before the start date of OL-OT (for subjects moving into OL-OT), or</li> <li>• before the SFE period start (for subjects moving into SFE period)</li> </ul>	Open-label CZP TEAE
	3. Start date is on or after the start date of OL-OT.  <b>Note:</b> Other treatment adverse events will not be summarized, but only presented in listings	Open-label other treatment TEAE
Subjects who enter the pre-SFE period	Start date is after the last double-blind or OL-CZP dose + 70 days and is before the SFE period start.	Pre-SFE AE
Subjects who enter the open-label SFE period	Start date is on or after the SFE period start and on or before the SFE period end + 70 days.	SFE TEAE

---

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