STATISTICAL ANALYSIS PLAN

AMENDMENT 2

Study: AS0005

Product: Certolizumab Pegol

MULTICENTER, OPEN-LABEL (PART A) FOLLOWED BY A RANDOMIZED, DOUBLE-BLIND, PARALLEL-GROUP, PLACEBO-CONTROLLED STUDY (PART B) TO EVALUATE MAINTENANCE OF REMISSION IN SUBJECTS WITH ACTIVE AXIAL SPONDYLOARTHRITIS (AXSPA) RECEIVING EITHER CERTOLIZUMAB PEGOL 200MG Q2W OR 200MG Q4W AS COMPARED TO PLACEBO

SAP/Amendment Number	Date
Final SAP	30 Mar 2017
SAP Amendment 1	23 April 2018
SAP Amendment 2	07 May 2019

TABLE OF CONTENTS

LIST OF ABBREVIATIONS 9
1 INTRODUCTION 13
2 PROTOCOL SUMMARY 13
2.1 Study objectives 13
2.1.1 Primary objectives 13
2.1.2 Secondary objectives 13
2.1.3 Other objectives 14
2.1.4 Pharmacokinetic, pharmacogenomic and immunological objectives 14
2.2 Study variables 14
2.2.1 Efficacy variables 14
2.2.1.1 Primary efficacy variable 14
2.2.1.2 Secondary efficacy variables 14
2.2.1.3 Other efficacy variables 15
2.2.2 Pharmacokinetic, pharmacogenomic and immunogenicity variables 19
2.2.2.1 Pharmacokinetic variables 19
2.2.2.2 Biomarkers 19
2.2.2.3 Pharmacogenomic variables 19
2.2.2.4 Immunogenicity variables 19
2.2.3 Safety variables 20
2.3 Study design and conduct 20
2.3.1 Part A 20
2.3.2 Part B 20
2.3.3 Escape treatment 21
2.3.4 Study duration per subject 21
2.3.5 Planned number of subjects and sites 22
2.4 Determination of sample size 22
3 DATA ANALYSIS CONSIDERATIONS 22
3.1 General presentation of summaries and analyses 22
3.2 General study level definitions 23
3.2.1 Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) 23
3.2.2 ASAS20, ASAS40, ASAS5/6 response, and ASAS Partial Remission 24
3.2.3 Bath Ankylosing Spondylitis Functional Index (BASFI) 25
3.2.4 MRI assessments 25
3.2.5 Bath Ankylosing Spondylitis Metrology Index (BASMI) 27
3.2.6 Total and nocturnal spinal pain NRS 28
3.2.7 Ankylosing Spondylitis Disease Activity Score (ASDAS) 28
3.2.8 Ankylosing Spondylitis Quality of Life (ASQoL) 30

3.2.9 Work Productivity Survey (WPS) 30
3.2.10 Patient's Global Assessment of Disease Activity (PtGADA, NRS) 31
3.2.11 Short-Form 36-item Health Survey (SF-36) 32
3.2.12 Maastricht Ankylosing Spondylitis Enthesitis Score (MASES) 32
3.2.13 Swollen and Tender Joint Counts (44) 32
3.2.14 Physician's Global Assessment of Disease Activity (PhGADA) 33
3.2.15 Spinal mobility 33
3.2.16 Health status (EQ-5D-3L) 33
3.2.17 Resource utilization 33
3.2.18 Extra-articular assessments 34
3.2.19 Fecal and serum calprotectin 34
3.2.20 Inflammatory Bowel Disease Questionnaire (IBD-Q) 34
3.2.21 Sustained Remission 36
3.2.22 Flare 36
3.2.22.1 Observed case 37
3.2.22.2 Alternative handling of premature withdrawal 37
3.2.22.3 Mixed model with repeated measures (MMRM) imputation of ASDAS 37
3.2.22.4 Referenced based multiple imputation of ASDAS 38
3.2.23 Time to flare 39
3.2.24 Subject age 41
3.2.25 Subject race 41
3.2.26 Anti-CZP antibody status 41
3.2.27 Modified New York criteria 41
3.2.28 Modified ASAS Classification Criteria 42
3.2.29 Past, prior and concomitant medications 42
3.3 Analysis time points 42
3.3.1 Relative day 44
3.3.2 End date of the Treatment Period 45
3.4 Definition of Baseline values 45
3.5 Protocol deviations 46
3.6 Analysis sets 46
3.6.1 Enrolled Set 46
3.6.2 Open-Label Set 46
3.6.3 Randomized Set 46
3.6.4 Safety Set 46
3.6.5 Safety Set Part B 46
3.6.6 Full Analysis Set 46
3.6.7 Per Protocol Set 46

3.6.8 Pharmacokinetic Set B 47
3.6.9 Flared Set 47
3.6.10 Escape Therapy Set 47
3.6.11 Part B Full-dose Set 47
3.6.12 General use of the Analysis Sets 47
3.7 Treatment assignment and treatment groups 47
3.8 Center pooling strategy 48
3.9 Coding dictionaries 48
3.10 Changes to protocol-defined analyses 48
4 STATISTICAL/ANALYTICAL ISSUES 49
4.1 Adjustments for covariates 49
4.2 Handling of dropouts or missing data 49
4.3 Interim analyses and data monitoring 53
4.4 Multicenter studies 54
4.5 Multiple comparisons/multiplicity 54
4.6 Use of an efficacy subset of subjects 54
4.7 Examination of subgroups 54
5 STUDY POPULATION CHARACTERISTICS 55
5.1 Subject disposition 55
5.2 Protocol deviations 55
6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS 56
6.1 Demographics 56
6.2 Other Baseline characteristics 56
6.3 Medical history and concomitant diseases 57
6.4 Past, prior and concomitant medications 58
7 MEASUREMENTS OF TREATMENT COMPLIANCE 59
8 EFFICACY ANALYSES 60
8.1 Primary efficacy variable 60
8.1.1 Primary analysis of the primary efficacy variable 60
8.1.2 Secondary analyses of the primary efficacy variable 61
8.1.3 Supportive and sensitivity analyses of the primary efficacy variable 61
8.2 Secondary efficacy variables 62
8.2.1 Part A 62
8.2.1.1 Percentage of subjects achieving sustained remission at Week 48 62
8.2.1.2 ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD and ASDAS-vHD) at Week 48 62
8.2.1.3 ASDAS clinical improvement (ASDAS-MI and ASDAS-CII) at Week 4862
8.2.2 Part B 62

8.2.2.1 Time to flare 62
8.2.2.2 Continuous secondary efficacy variables 63
8.2.2.3 Binary secondary efficacy variables 63
8.2.3 Subjects experiencing a flare in Part B 64
8.2.3.1 Continuous secondary efficacy variables 64
8.2.3.2 Binary secondary efficacy variables 64
8.3 Other efficacy variables 64
8.3.1 Part A 64
8.3.1.1 ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD and ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI) 64
8.3.1.2 ASAS20, ASAS40, ASAS5/6, and ASAS PR response 64
8.3.1.3 ASDAS, BASDAI, and BASMI 65
8.3.1.4 SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores 65
8.3.1.5 Individual ASAS core components 65
8.3.1.6 PhGADA 65
8.3.1.7 Nocturnal spinal pain (NRS) 65
8.3.1.8 Fatigue (NRS) (from BASDAI) 65
8.3.1.9 BASDAI50 response 66
8.3.1.10 C-reactive protein (CRP) 66
8.3.1.11 Ankylosing Spondylitis Quality of Life (ASQoL) 66
8.3.1.12 Work Productivity Survey (WPS) 66
8.3.1.13 EuroQoL Health Status Questionnaire 5 dimensions 3 levels (EQ-5D-3L) 66
8.3.1.14 Maastricht Ankylosis Spondylitis Enthesitis Score (MASES) 66
8.3.1.15 Swollen and tender joint counts (44 joint count) 66
8.3.1.16 Spinal mobility 67
8.3.1.17 Number of uveitis flares, IBD exacerbations and psoriasis exacerbations 67
8.3.1.18 Fecal calprotectin 67
8.3.1.19 Serum calprotectin 67
8.3.1.20 Inflammatory Bowel Disease Questionnaire (IBD-Q) 67
8.3.1.21 Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP 68
8.3.1.22 Short-Form 36-Item Health Survey (SF-36) 68
8.3.1.23 Resources utilization 68
8.3.2 Part B 68
8.3.2.1 ASDAS clinical improvement (ASDAS-CII, ASDAS-MI) 68
8.3.2.2 ASAS20, ASAS40, ASAS5/6, and ASAS PR response 69
8.3.2.3 ASDAS, BASDAI, BASFI and BASMI 69
8.3.2.4 Individual ASAS core components 69
8.3.2.5 Fatigue NRS 70

8.3.2.6 BASDAI50 response 70
8.3.2.7 ASQoL 70
8.3.2.8 Work Productivity Survey (WPS) 70
8.3.2.9 Number of uveitis flares, IBD exacerbations and psoriasis exacerbations 71
8.3.2.10 Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP 71
8.3.2.11 Short-Form 36-Item Health Survey (SF-36) 71
8.3.2.12 Descriptive summaries of other efficacy variables 72
8.3.3 Subjects experiencing a flare in Part B 73
9 PHARMACOKINETICS AND PHARMACODYNAMICS 74
9.1 Pharmacokinetics 74
9.2 Immunogenicity 75
10 SAFETY ANALYSES 77
10.1 Extent of exposure 77
10.1.1 Part A 77
10.1.2 Part B 78
10.1.2.1 All subjects in SSB – Prior to Flare 78
10.1.2.2 After flare 79
10.1.2.3 Full-dose CZP in Part B 79
10.1.3 Part A and Part B full-dose CZP therapy 80
10.2 Adverse events 82
10.3 Clinical laboratory evaluations 90
10.4 Vital signs, physical findings, and other observations related to safety 94
10.4.1 Vital signs 94
10.4.2 Other safety variables 94
10.4.2.1 Pregnancy testing 94
10.4.2.2 Physical assessments 94
10.4.2.3 Tuberculosis assessments 95
11 REFERENCES 96
12 APPENDICES 97
12.1 Rules for medical resource utilization 97
12.1.1 Healthcare Resource Utilization 97
12.1.2 Counting rules 97
12.1.2.1 In-patient hospitalization and emergency room visits 97
12.1.2.2 Healthcare provider consultations not foreseen by the protocol 98
12.1.2.3 Concomitant medical procedures 98
12.2 Compliance ratio calculation 98
12.3 SAP Amendment 1 100

12.4 SAP Amendment 2 111

LIST OF TABLES

Table 3.1: BASMI linear definition 28
Table 3.2: Tender and swollen joint scoring 32
Table 3.3: Flare Definition and Time to Flare 41
Table 3.4: List of countries in each region 48
Table 10.1: Definitions of Markedly Abnormal Hematology Values 92
Table 10.2: Definitions of Markedly Abnormal Biochemistry Values 92
Table 12.1: Example 99

LIST OF ABBREVIATIONS

ADAb	anti-CZP antibody
ABS	absolute value
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
AP	alkaline phosphatase
AS	ankylosing spondylitis
ASAS	Assessment of SpondyloArthritis international Society
ASAS20, 40	Assessment of SpondyloArthritis international Society 20%, 40% response criteria
ASAS5/6	ASAS 5 out of 6 response criteria
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 activity
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 score for activity
AST	aspartate aminotransferase
ATC	anatomical therapeutic chemical
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
BLQ	below the limit of quantification
CD	Crohn's disease

CI	confidence interval
CPK	creatine phosphokinase
CRP	C-reactive protein
CSR	Clinical Study Report
CV	coefficient of variation
CZP	certolizumab pegol
DM	data management
DMARD	disease-modifying antirheumatic drug
DOB	date of birth
EQ-5D-3L	EuroQol Health Status Questionnaire 5 dimensions 3 levels
EAIR	exposure adjusted incidence rate
EAER	exposure adjusted event rate
eCRF	electronic Case Report Form
ES	Enrolled Set
ETS	Escape Therapy Set
FAS	Full Analysis Set
FS	Flared Set
H	high
HIV	human immunodeficiency virus
HLA-B27	human leukocyte antigen B27
HLT	High Level Term
HRQoL	Health-related Quality of Life
IBD	inflammatory bowel disease
IBD-Q	Inflammatory Bowel Disease Questionnaire
ICH	International Council for Harmonisation
IGRA	interferon-gamma release assay
IXRS	interactive voice or web response system
L	low
LLOQ	lower limit of quantification
LOCF	last observation carried forward
MAR	missing at random
MASES	Maastricht Ankylosis Spondylitis Enthesitis Score
MCMC	Markov chain Monte Carlo
MCS	Mental Component Summary

MedDRA®	Medical Dictionary for Regulatory Activities
MH	marked abnormal high
ML	marked abnormal low
MMRM	Mixed model with repeated measures
mNY	modified New York
MRI	magnetic resonance imaging
MSR	minimum significant ratio
N	normal
NA	not applicable
NAb	neutralizing antibody
nr-axSpA	non-radiographic axial spondyloarthritis
NRI	non-responder imputation
NRS	numeric rating scale
NSAID	non-steroidal anti-inflammatory drug
OLS	Open-Label Set
PCS	Physical Component Summary
PDS	Protocol Deviation Specification
PhGADA	Physician's Global Assessment of Disease Activity
PK	pharmacokinetics
PKSA	Pharmacokinetic Set A
PKSB	Pharmacokinetic Set B
PBO	placebo
PPS	Per Protocol Set
PR	partial remission
PT	preferred term
PtGADA	Patient's Global Assessment of Disease Activity
Q2W	every 2 weeks
Q4W	every 4 weeks
r-axSpA	radiographic axial spondyloarthritis
RCTC	Rheumatology Common Toxicity Criteria
RS	Randomized Set
SAP	Statistical Analysis Plan
SD	standard deviation
SDTM	Study Data Tabulation Model

SF-36	Short-Form 36-Item Health Survey
SFU	Safety Follow Up
SIJ	Sacroiliac Joint
SMQ	standard MedDRA® query
SOC	system organ class
SPARCC	Sacroiliac SpondyloArthritis Research Consortium of Canada
SS	Safety Set
SSB	Safety Set Part B
STIR	short-tau-inversion recovery
TB	tuberculosis
TE	treatment-emergent
TNF	tumor necrosis factor
UK	United Kingdom
ULN	upper limit of normal
VAS	visual analog scale
VU	vertebral units
WHO-DD	World Health Organization Drug Dictionary
WPS	Work Productivity Survey

1 INTRODUCTION

This Statistical Analysis Plan (SAP) is designed to support a Clinical Study Report (CSR). This SAP complies with International Council for Harmonisation (ICH) guidelines ICH E3 (Structure and Content of Clinical Study Reports, 1995), ICH E9 (Statistical Principles for Clinical Trials, 1998) and ICH E10 (Choice of Control Group in Clinical Trials, 2000) and is based upon the following protocol and amendments:

- Final Protocol dated 27 Mar 2015
- Protocol Amendment 0.1 (Taiwan) dated 13 Aug 2015
- Protocol Amendment 0.2 United Kingdom (UK) dated 31 Jul 2015
- Protocol Amendment 1 dated 24 Nov 2015
- Protocol Amendment 1.1 (Taiwan) dated 18 Dec 2015
- Protocol Amendment 1.2 (UK) dated 18 Dec 2015
- Protocol Amendment 2 dated 24 Jan 2018
- Protocol Amendment 2.1 (Taiwan) dated 27 Feb 18
- Protocol Amendment 2.2 (UK) dated 27 Feb 18

This SAP does not cover the analysis of genetics, genomics and proteomics data.

2 PROTOCOL SUMMARY

2.1 Study objectives

2.1.1 Primary objectives

The primary objective of the study is to evaluate the percentage of subjects who do not experience a flare on certolizumab pegol (CZP) 200mg every 2 weeks (Q2W; full-dose) or 200mg every 4 weeks (Q4W; half-dose) as compared to placebo (PBO) (CZP withdrawal) during Part B.

2.1.2 Secondary objectives

The secondary objectives are to:

- Evaluate the percentage of subjects achieving sustained remission at the end of Part A
- For subjects randomized into Part B:
  - Evaluate the time to flare and other measures of signs and symptoms
  - Compare the percentage of subjects who do not experience a flare between CZP full-dose and half-dose
  - Evaluate the efficacy of re-initiation of treatment with the CZP full-dose in subjects who experience a flare following a withdrawal or dose reduction of CZP
- Assess safety and tolerability of CZP
- Evaluate inflammatory changes over time as assessed by magnetic resonance imaging (MRI)

2.1.3 Other objectives

The other objectives are to evaluate:

- Physical function
- Signs and symptoms of the disease
  - Morning stiffness
  - Fatigue
  - Extra-articular manifestations of axial spondyloarthritis (axSpA)
- Health-related Quality of Life (HRQoL)
- Work and household productivity
- Subject's health status
- Direct medical resource utilization
- Correlation of fecal calprotectin, serum calprotectin, Inflammatory Bowel Disease Questionnaire (IBD-Q) and C-reactive protein (CRP)

2.1.4 Pharmacokinetic, pharmacogenomic and immunological objectives

Pharmacokinetic and immunological objectives are to follow the CZP concentrations throughout the study, as well as the evolution of immunogenicity, respectively. Other pharmacogenomic objectives are:

- Exploratory biomarkers and cytokines
- Possible genetics, genomics, and proteomics

2.2 Study variables

2.2.1 Efficacy variables

2.2.1.1 Primary efficacy variable

The primary efficacy variable is the percentage of subjects in Part B who do not experience a flare.

2.2.1.2 Secondary efficacy variables

2.2.1.2.1 Secondary efficacy variables for subjects entering Part A

- Percentage of subjects achieving sustained remission at Week 48
- Ankylosing Spondylitis Disease Activity Score (ASDAS) disease activity (ASDAS-Inactive Disease [ASDAS-ID], ASDAS-Low Disease activity [ASDAS-LD] (formally called ASDAS Moderate Diasease), ASDAS-High Disease activity [ASDAS-HD], and ASDAS-very High Disease activity [ASDAS-vHD]) and clinical improvement (ASDAS-Clinically Important Improvement [ASDAS-CII], ASDAS-Major Improvement [ASDAS-MI]) at Week 48

2.2.1.2.2 Secondary efficacy variables for subjects entering Part B

- Time to flare

- ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD, and ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI) at Week 96
- Assessment of SpondyloArthritis International Society (ASAS) response criteria ASAS20, ASAS40, ASAS5/6, and ASAS partial remission (PR) response at Week 96
- Change from Part A Baseline in ASDAS, Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Functional Index (BASFI), and Bath Ankylosing Spondylitis Metrology Index (BASMI) at Week 96
- BASDAI50 response at Week 96
- Change from Part A Baseline in Sacroiliac Joint (SIJ) Sacroiliac SpondyloArthritis Research Consortium of Canada (SPARCC) and spine ankylosing spondylitis spine MRI score for activity (ASspiMRI-a) in the Berlin modification scores at Week 96

2.2.1.2.3 Secondary efficacy variables for subjects who experience a flare in Part B

These variables will be evaluated at Escape Week 12.

- ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD, and ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI)
- ASAS20, ASAS40, ASAS5/6, and ASAS PR response
- Change from Part B Baseline in ASDAS, BASDAI, BASFI and BASMI
- Change from Part B Baseline in SIJ SPARCC and spine ASspiMRI-a in Berlin modification scores

2.2.1.3 Other efficacy variables

2.2.1.3.1 Other efficacy variables for subjects entering Part A

The following variables will be evaluated at all scheduled study visits where the assessment is performed through Week 48.

- ASAS20, ASAS40, ASAS5/6, and ASAS PR response
- BASDAI50 response
- Change from Part A Baseline in ASDAS, BASDAI and BASMI
- ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD, and ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI)
- Change from Part A Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores
- Change from Part A Baseline in all individual ASAS core components, ie:
  - Patient's Global Assessment of Disease Activity (PtGADA)
  - Total spinal pain (Numeric Rating Scale [NRS])
  - BASFI

  - Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
- Change from Part A Baseline in nocturnal spinal pain (NRS)
- Change from Part A Baseline in spinal mobility, as assessed by occiput to wall distance
- Change from Part A Baseline in spinal mobility, as assessed by chest expansion
- Change from Part A Baseline in Physician's Global Assessment of Disease Activity (PhGADA)
- Change from Part A Baseline in Fatigue (NRS) (from BASDAI)
- Change from Part A Baseline in CRP
- Change from Part A Baseline in Ankylosing Spondylitis Quality of Life (ASQoL)
- Change from Part A Baseline in Work Productivity Survey (WPS)
- Health status as assessed by the EuroQoL Health Status Questionnaire 5 dimensions 3 levels (EQ-5D-3L): domains, visual analog scale (VAS) actual score and change from Part A Baseline in VAS score
- Change from Part A Baseline in Maastricht Ankylosis Spondylitis Enthesitis Score (MASES)
- Change from Part A Baseline in swollen and tender joint counts (44 joint count)
- Number of uveitis flares
- Number of inflammatory bowel disease (IBD) exacerbations
- Change from Part A Baseline in fecal calprotectin
- Change from Part A Baseline in serum calprotectin
- Change from Part A Baseline in the IBD-Q
- Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP
- Number of psoriasis exacerbations (in subjects with concurrent psoriasis)
- Change from Part A Baseline in all Short-Form 36-Item Health Survey (SF-36) domains, SF-36 Physical Component Summary (PCS), and SF-36 Mental Component Summary (MCS)
- Resources utilization: concomitant medical procedures, healthcare provider consultations not foreseen by the protocol, in-patient hospitalizations and emergency room visits

2.2.1.3.2 Other efficacy variables for subjects entering Part B

The following variables will be evaluated at all scheduled study visits where the assessment is performed following the first administration of study medication in Part B.

- ASAS20, ASAS40, ASAS5/6, and ASAS PR response using Part A Baseline as a reference
- BASDAI50 response using Part A Baseline as a reference

- ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD, and ASDAS-vHD) and ASDAS clinical improvement (ASDAS-CII, ASDAS-MI) using Part A Baseline as a reference
- Change from Part A Baseline in ASDAS, BASDAI and BASMI
- Change from Part B Baseline in ASDAS, BASDAI, BASFI and BASMI
- Change from Part A Baseline and change from Part B Baseline in all individual ASAS core components, ie:
  - PtGADA
  - Total spinal pain (NRS)
  - BASFI
  - Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
- Change from Part A Baseline and change from Part B Baseline in nocturnal spinal pain (NRS)
- Change from Part A Baseline and change from Part B Baseline in spinal mobility, as assessed by occiput to wall distance
- Change from Part A Baseline and change from Part B Baseline in spinal mobility, as assessed by chest expansion
- Change from Part A Baseline and change from Part B Baseline in PhGADA
- Change from Part A Baseline and change from Part B Baseline in Fatigue (NRS) (from BASDAI)
- Change from Part A Baseline and change from Part B Baseline in CRP
- Change from Part A Baseline and change from Part B Baseline in ASQoL
- Change from Part A Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores
- Change from Part B Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores
- Change from Part A Baseline and change from Part B Baseline in WPS
- Health status as assessed by the EQ-5D-3L: domains, VAS actual score, change from Part A Baseline in VAS score and change from Part B Baseline in VAS score
- Change from Part A Baseline and change from Part B Baseline in MASES
- Change from Part A Baseline and change from Part B Baseline in swollen and tender joint counts (44 joint count)
- Number of uveitis flares
- Number of IBD exacerbations
- Change from Part A Baseline and change from Part B Baseline in fecal calprotectin

- Change from Part A Baseline and change from Part B Baseline in serum calprotectin
- Change from Part A Baseline and change from Part B Baseline in the IBD-Q
- Number of psoriasis exacerbations (in subjects with concurrent psoriasis)
- Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP
- Number of psoriasis exacerbations (in subjects with concurrent psoriasis)
- Change from Part A Baseline and change from Part B Baseline in all SF-36 domains, SF-36 PCS, and SF-36 MCS
- Resources utilization: concomitant medical procedures, healthcare provider consultations not foreseen by the protocol, in-patient hospitalizations and emergency room visits

2.2.1.3.3 Other efficacy variables for subjects who experience a flare in Part B

The following variables will be evaluated at all scheduled study visits for subjects experiencing a flare in Part B where the assessment is performed following escape week 0.

- ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD, and ASDAS-vHD) andASDAS clinical improvement (ASDAS-CII, ASDAS-MD) using Flare Baseline as reference
- ASAS20, ASAS40, ASAS5/6, and ASAS PR response using Flare Baseline as reference
- Change from Part B Baseline in ASDAS, BASDAI, BASFI and BASMI
- Change from Flare Baseline in ASDAS, BASDAI, BASFI and BASMI
- Change from Part A Baseline, change from Part B Baseline and change from Flare Baseline in all individual ASAS core components, ie:
  - PtGADA
  - Total spinal pain (NRS)
  - BASFI
  - Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
- Change from Part B Baseline and change from Flare Baseline in nocturnal spinal pain (NRS)
- Change from Part B Baseline and change from Flare Baseline in spinal mobility, as assessed by occiput to wall distance
- Change from Part B Baseline and change from Flare Baseline in spinal mobility, as assessed by chest expansion
- Change from Part B Baseline and change from Flare Baseline in PhGADA
- Change from Part B Baseline and change from Flare Baseline in Fatigue (NRS) (from BASDAI)
- BASDAI50 response using Flare Baseline as reference
- Change from Part B Baseline and change from Flare Baseline in CRP

- Change from Part B Baseline and change from Flare Baseline in ASQoL
- Change from Part A Baseline in SIJ SPARCC and Spine ASspiMRI-a in Berlin modification scores
- Change from Part B Baseline in SIJ SPARCC and Spine ASspiMRI-a in Berlin modification scores
- Change from Flare Baseline in SIJ SPARCC and Spine ASspiMRI-a in Berlin modification scores
- Change from Part B Baseline and change from Flare Baseline in WPS
- Health status as assessed by the EQ-5D-3L: domains, VAS actual score, and change from Part B Baseline and change from Flare Baseline in VAS score
- Change from Part B Baseline and change from Flare Baseline in MASES
- Number of uveitis flares
- Number of IBD exacerbations
- Change from Part B Baseline and change from Flare Baseline in fecal calprotectin
- Change from Part B Baseline and change from Flare Baseline in serum calprotectin
- Change from Part B Baseline and change from Flare Baseline in the IBD-Q
- Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP
- Number of psoriasis exacerbations (in subjects with concurrent psoriasis)
- Change from Part B Baseline and change from Flare Baseline in all SF-36 domains, SF-36 PCS, and SF-36 MCS
- Resources utilization: concomitant medical procedures, healthcare provider consultations not foreseen by the protocol, in-patient hospitalizations and emergency room visits

2.2.2 Pharmacokinetic, pharmacogenomic and immunogenicity variables

2.2.2.1 Pharmacokinetic variables

Plasma concentrations of CZP will be measured at Part A Baseline and subsequent time points as described in Table 5.1 and Table 5.2 of the protocol only for subjects randomized into Part B (see Section 3.10). These plasma samples may be used additionally for analyses of CZP and its constituent moieties using alternative methods.

2.2.2.2 Biomarkers

Any analysis for the biomarker will be described in a separate document.

2.2.2.3 Pharmacogenomic variables

Any analysis for the pharmacogenomic will be described in a separate document.

2.2.2.4 Immunogenicity variables

Anti-CZP antibody (ADAb) levels will be assessed at Part A Baseline and subsequent time points as described in the protocol.

Determination of ADAb will be done using a validated screening, confirmation, and titration ADAb bridging assay, with potential further characterization by a neutralizing antibody (NAb) assay. The NAb will be not part of the CSR. The immunogenicity samples will be processed according to dedicated Bioanalytical Analysis plans.

2.2.3 Safety variables

Safety variables to be assessed are adverse events (AEs), physical examination, vital signs, chest x-ray, interferon-gamma release assay (IGRA) test for tuberculosis, and measurements of laboratory parameters (hematology, biochemistry, and urinalysis). The safety variables will be stated or listed as described within the appropriate Sections.

2.3 Study design and conduct

Study AS0005 is a multicenter, open-label (Part A) followed by a randomized, double-blind, parallel-group, PBO-controlled (Part B) clinical study to evaluate the efficacy, safety, pharmacokinetics (PK), and immunogenicity of CZP in adult subjects with axSpA in sustained remission who continued either on full-dose treatment (CZP 200mg Q2W), on a dose reduction (CZP 200mg Q4W), or withdrawal of CZP treatment. The study includes 2 parts, A and B, as follows.

2.3.1 Part A

Part A consists of two study periods.

Period 1 (Screening Period): Up to 5 Weeks before Baseline

Period 2 (Open-Label Period): Week 0 (Baseline) to Week 48

Within Period 2, eligible subjects will receive 3 loading doses of CZP 400mg at Weeks 0, 2, and 4 followed by CZP 200mg Q2W.

All subjects who have not achieved sustained remission at the end of Period 2 (Week 48) will be discontinued from the study and treated at the discretion of their physician. Those subjects who discontinue prematurely from the study for any reason will be treated at the discretion of their physician.

2.3.2 Part B

Period 3 (Double-Blind Period): Week 48 to Week 96, PBO-controlled

Subjects must meet the sustained remission criteria in order to be randomized to Part B. These subjects in sustained remission will be randomized in a 1:1:1 ratio to one of the following treatment arms:

- CZP administered at a dose of 200mg Q2W (full-dose)
- CZP administered at a dose of 200mg Q4W (half-dose)
- PBO

All subjects who discontinue prematurely from the study for any reason during the Double-Blind Period will be treated at the discretion of the Investigator.

The final study assessments will be performed at Week 96. Subjects experiencing a flare during the Double-Blind Period will receive escape-treatment as defined in Section 2.3.3.

Period 4 (Safety Follow Up Period): 10 weeks after the last dose of study medication

All subjects, including those withdrawn from study treatment, will have a Safety Follow Up (SFU) Visit 10 weeks after their last dose of study medication. Subjects in UK will be followed up additionally for 12 weeks after the SFU Visit, by phone call for occurrence of any clinically significant infections, including tuberculosis.

2.3.3 Escape treatment

Subjects experiencing a flare in Part B will be invited for an on-site visit to receive the next planned dose of investigational product (escape treatment). This visit will take place 2 weeks after the last assessment that resulted in meeting the flare criteria; or, at the discretion of the Investigator, at an earlier time point. As the start of escape treatment may vary from subject to subject, it will be referred to as "Escape Week 0". Subjects will remain on the escape treatment for at least 12 weeks or until Week 96 of the regular visit schedule as described in Table 5.1 in the protocol, whichever is longer.

The assessment-schedule for subjects experiencing a flare in Part B will start 2 weeks after the flare condition is fulfilled; or, at the discretion of the Investigator, at an earlier time point and will end 2 weeks before Week 96 after Baseline of the regular assessment schedule. If the flare occurs late in Part B (at or after Week 82), the subject must receive the escape treatment for 12 weeks (eg, starting at Week 84 until Week 96). The final assessments at study end (as laid down for Week 96) must be performed 2 weeks after the last study dose administration accordingly.

Subjects receiving PBO or half-dose CZP (200mg Q4W) who experience a flare during the Double-Blind Period will escape to full-dose treatment (CZP 200mg Q2W) until the end of that period or for at least 12 weeks, whichever is longer. The escapers need to come back to the site at the time point of the escape (Week 0), 2 and 4 weeks after escape (Weeks 2 and 4), and every 12 weeks thereafter until the originally scheduled Week 96.

Subjects escaping from PBO to full-dose CZP (200mg Q2W) will receive a loading dose of CZP (400mg Q2W) at the 3 consecutive visits after the flare. Subjects escaping from half-dose CZP to full-dose CZP will receive CZP 200mg Q2W at the 3 consecutive visits after the flare. In order to maintain the blind, subjects escaping from half-dose CZP to full-dose CZP will receive a PBO-administration and a CZP 200mg administration at these 3 visits.

Subjects randomized to the CZP full-dose (200mg Q2W) treatment during the Double-Blind Period who qualify for escape will remain on their current treatment allocation. As with subjects escaping from half-dose CZP to full-dose CZP, these subjects will receive a PBO and CZP 200mg injection at the 3 visits after the flare in order to maintain the blind to the randomized treatment.

After the loading dose, all escaped subjects will continue full-dose CZP treatment in an open-label fashion.

2.3.4 Study duration per subject

For each subject, there is a planned duration of up to 109 weeks, as follows:

- Up to 5 weeks of Screening Period
- 48 weeks in the Open-Label Period
- 48 weeks in the Double-Blind Period

- An SFU Visit 10 weeks after last dose administration (Period 4).

Of note, the last dosing in the Double-Blind Period occurs at Week 94 so that the SFU of 10 weeks already includes 2 weeks from the Double-Blind Period hence leading to a duration of 109 weeks.

Subjects in UK will be contacted by a telephone call 22 weeks after the last dose of their study medication hence leading to a duration of 121 weeks.

For subjects who experience a flare during the Double-Blind Period, the study duration might be prolonged.

2.3.5 Planned number of subjects and sites

Approximately 1250 subjects will be screened in order to enroll 750 subjects into Part A, where 210 subjects are expected to meet the sustained remission criteria and be eligible for randomization into Part B. The ASDAS scores will be closely monitored during Part A of the study to project the percentage of enrolled subjects likely to achieve sustained remission at the end of Part A. The enrollment will be adjusted accordingly in order to achieve the required number of 210 subjects in sustained clinical remission qualifying for Part B.

The end of the study is defined as the date of the last visit (SFU) of the last subject in the study. It is planned to enroll the subjects at approximately 95 sites.

The study will be conducted in North America, Western Europe, Eastern Europe and Asia.

2.4 Determination of sample size

It is anticipated that approximately 210 subjects will be randomized in a 1:1:1 ratio to the CZP 200mg Q2W, CZP 200mg Q4W, and PBO treatment groups. The primary efficacy analysis is based on the proportion of subjects who do not experience a flare during Part B of the study. The proportion of subjects who do not experience a flare during Part B is assumed to be 80%, 75%, and 45% for the CZP 200mg Q2W, CZP 200mg Q4W, and PBO treatment groups, respectively. Given these assumptions, a sample size of 70 subjects per group provides 98% power to detect a significant difference between the CZP 200mg Q2W dose and PBO and 94% power to detect a significant difference between the CZP 200mg Q4W dose and PBO using a 2-sided significance level of 0.05.

Given that subjects enter Part B only if they have achieved sustained remission in Part A, it is necessary to estimate the number of subjects needed to enroll in Part A. It is assumed that approximately 28% of subjects enrolled in Part A will achieve sustained remission and qualify for Part B. This means that about 750 subjects will be enrolled in Part A in order to have 210 subjects who are randomized into Part B.

3 DATA ANALYSIS CONSIDERATIONS

3.1 General presentation of summaries and analyses

Statistical analysis and generation of tables, figures, and listings will be performed using SAS® Version 9.3 or higher.

Frequency tables (frequency counts and percentages) will be presented for categorical data. If there are missing values, either a missing category will be included in the display or the number of non-missing results will be used for calculations.

In case imputation was performed, summary statistics will not utilize the 'n' (number of available measurements). In addition, for the primary variables and selected outputs the 95% confidence interval (CI) based on the Exact Binomial distribution will be presented.

In general percentages will be calculated based on the utilized analysis set, however in case only subsets are affected, the N of the subset will be used as denominator.

Unless otherwise noted, all percentages will be displayed to 1 decimal place. No percentage will be displayed for zero counts, and no decimal will be presented when the percentage is 100%.

Continuous variables will be summarized with n (number of available measurements), arithmetic mean, standard deviation (SD), median, minimum, and maximum (with 25th and 75th percentiles where specified). For CZP plasma concentration levels and CRP observed values and ratio of post-Baseline over Baseline values the geometric mean and the corresponding coefficient of variation (CV) will be presented. The CV will be calculated as

$$CV = \sqrt{e^{std\ ln^2} - 1} * 100$$

where "std ln" represents the standard deviation of the natural logarithm transformed parameter.

The mean, SD, and median will be displayed to 1 more decimal place than collected in the electronic Case Report Form (eCRF). Minimum and maximum values will be displayed to the same level of precision as collected in the eCRF. Unless otherwise specified, all statistical hypothesis tests will be 2-sided and the level of significance for all tests will be 0.05. P-values will be displayed to 3 decimal places.

The study design has three different periods: Part A, Part B and Escape Part. The outputs will be presented for each period. Summaries for Part A will be presented by axSpA subpopulation. Summaries for Part B will be presented by treatment group excluding any post-flare assessments in the subset of subjects experiencing a flare in Part B and start escape therapy. For the subset of subjects experiencing a flare in Part B and start escape therapy, separate summaries will be produced for the Escape Part.

3.2 General study level definitions

3.2.1 Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)

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.

The BASDAI is calculated as follows:

$$BASDAI = \frac{Q1 + Q2 + Q3 + Q4 + \frac{Q5 + Q6}{2}}{5}$$

When calculating observed case value, following rules will be applied:

- [redacted]

[redacted]

- 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. No imputation will be performed.

When performing last observation carried forward (LOCF) imputation, following rules will be applied:

- If more than one but not all components are missing at a visit, the previous value for the missing components will be carried forward to replace the missing components, assuming the previous values were from a post-baseline assessment and used to calculate the BASDAI score. If there are no previous post-baseline values available for any of the missing components to carry forward, the previous visit's total BASDAI score will be carried forward, assuming the previous visit was a post-baseline assessment. Otherwise the LOCF score is missing at that visit.
- If all components are missing at a visit, the previous visit's total BASDAI score will be carried forward, assuming the previous visit was a post-baseline assessment. Otherwise the LOCF score is missing at that visit.

The response variable BASDAI50 will be defined as an improvement (reduction) of at least 50% in the BASDAI compared to Baseline.

Fatigue item of the BASDAI

Fatigue as a major symptom of ankylosing spondylitis (AS) can effectively be measured with single item questions such as the BASDAI item. This item has shown moderate to good reliability and responsiveness. The fatigue item score will range from 0 to 10.

3.2.2 ASAS20, ASAS40, ASAS5/6 response, and ASAS Partial Remission

The ASAS20 is defined as an improvement (decrease) of ≥20% and absolute improvement (decrease) of ≥1 unit on a 0 to 10 NRS in at least 3 out of 4 domains as follows:

- PtGADA (see Section 3.2.10)
- Pain assessment (total spinal pain NRS score)
- Function (represented by BASFI, Section 3.2.3)
- Inflammation (the mean of the BASDAI questions 5 and 6, [see Section 3.2.1] concerning morning stiffness intensity and duration)

and absence of deterioration in the potential remaining domain (deterioration is defined as a relative worsening (increase) of ≥20% and absolute worsening (increase) of ≥1 unit.

The ASAS40 is defined as an improvement (decrease) of ≥40% and absolute improvement (decrease) ≥2 units on a 0 to 10 NRS in at least 3 out of 4 domains and no worsening (increase) at all in the remaining domain.

The ASAS5/6 response is defined as an improvement of ≥20% in at least 5 out of 6 domains comprised of the 4 domains above plus additional domains of spinal mobility (ie, BASMI, lateral lumbar flexion), and CRP. The lateral lumbar flexion value is taken as the mean of left and right assessment values and improvement is a ≥20% increase. Improvement in CRP is a ≥20% decrease. If the CRP value is below the lower limit of quantification (LLOQ), then it will be imputed as the midpoint between 0 and the LLOQ.

The ASAS PR response is defined as a score of ≤2 units on a 0 to 10 NRS in each of the 4 domains listed above for ASAS20.

When calculating observed case response status for ASAS20, ASAS40, ASAS5/6 and ASAS PR, following rules will be applied:

- If 1 of the 2 morning stiffness measurements (ie, questions 5 and 6 of BASDAI questionnaire) is missing, the other one will be used for the inflammation calculation.
- If one or more components are missing at a visit, the response status will be treated as missing at that visit. No imputation will be performed. Subjects will not be handled as non-responder.

When performing non-responder imputation (NRI) for ASAS20, ASAS40, ASAS5/6 and ASAS PR, following rules will be applied:

- If some but not all components are missing at a visit, the previous value for the missing components will be carried forward to replace the missing values, assuming the previous values were from a post-baseline assessment. The response status will then be assessed. If there is no previous post-baseline value available to carry forward for any of the missing components, a 0 unit improvement value and 0% improvement value will be imputed for the missing components. The response status will then be assessed.
- If all components are missing at a visit, then the subject will be considered a non-responder at that visit.

3.2.3 Bath Ankylosing Spondylitis Functional Index (BASFI)

The BASFI comprises 10 items relating to the past week. The NRS version will be used for the answering options of each item on a scale of 0 (“Easy”) to 10 (“Impossible”). The BASFI is the mean of the 10 scores such that the total score ranges from 0 to10, with lower scores indicating better physical function.

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.

3.2.4 MRI assessments

Magnetic resonance imaging of the spine and sacroiliac joints will be performed at Screening, Week 48, Week 96, or Withdrawal Visit if the previous MRI was performed more than 12 weeks prior to withdrawal. In the Escape Part it will be also performed at Escape Week 0 and 12. MRI will be assessed centrally and scoring will be done according to the imaging charter by independent readers, who are blinded to both the order of the scans and to the treatment group,

using a previously reported scoring system. A single reader assessment will be performed for Week 0 and Week 48 in the MRI scores for all subjects that entered Part A and have at least one post-baseline scan. A double reader assessment will be performed for subjects entering Part B. The analysis will use the average of the scores from the 2 independent reviewers. If the two readers disagree a third reviewer who is not one of the two reviewers will perform adjudication by providing a third independent review. Whenever an adjudication is present the average score across all 3 reviewers will be used for analysis. Two different MRI scores will be used for analysis (SPARCC, Berlin modification of the ASspiMRI-a).

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 SIJ is divided into 4 quadrants and will be assessed for left and right: upper iliac, lower iliac, upper sacral, and lower sacral. The presence of increased signal on STIR in each of these 4 quadrants are 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 SIJ 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) or 24 VUs if a transitional vertebra is present. If a transitional vertebra is present, L5/L6 and L6/SI will be assessed instead of L5/SI. 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. This score will increase if there are transitional vertebrae.

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

- If all scores are not applicable (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.

For MRI the relative day after scheduled MRI visit will be calculated.

In case of out of window images/scores regarding the visit date, the following rules will apply:

- MRI assessments assigned to Baseline:
  - If Baseline MRI was performed from Screening up to 2 weeks after Baseline it will be assigned to Baseline, if beyond it will not be used.
- MRI assessments assigned to Week 48:
  - If Week 48 MRI was performed ±2 weeks around the Week 48 Visit, it will be assigned to Week 48, otherwise Week 48 will be treated as missing.
- MRI assessments assigned to Week 96:
  - If Week 96 MRI was performed ±2 weeks around the Week 96 Visit, it will be assigned to Week 96, otherwise Week 96 will be treated as missing.

In subjects experiencing a flare in Part B, following rules will apply:

- MRI assessments assigned to Escape Week 0:
  - If Escape Week 0 MRI was performed ±2 weeks around the Escape Week 0 Visit, it will be assigned to Escape Week 0, otherwise Escape Week 0 will be treated as missing.
- MRI assessments assigned to Escape Week 12:
  - If Escape Week 12 MRI was performed ±2 weeks around the Escape Week 12 Visit, it will be assigned to Escape Week 12, otherwise Escape Week 12 will be treated as missing.

In case a MRI is performed at a premature Withdrawal Visit, it will be assigned to what would have been the next scheduled visit with a MRI assessment. The MRI will be used for summaries if it was performed ±2 weeks around the scheduled visit it is assigned to.

3.2.5 Bath Ankylosing Spondylitis Metrology Index (BASMI)

The BASMI characterizes the spinal mobility of subjects with AS. The BASMI is a disease-specific measure consisting of 5 clinical measures to reflect subject axial status: cervical rotation; tragus to wall distance; lateral spinal flexion; modified Schober test; intermalleolar distance. 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.

Table 3.1: BASMI linear definition

S = (21.1 cm –A) / 2.1 cm	For the lateral lumbar spine flexion (mean right/left)
S = (A – 8 cm) / 3 cm	For the tragus-to-wall distance (mean right/left)
S = (7.4 cm – A) /0.7 cm	For the lumbar flexion (modified Schober)
S = (124.5 cm – A) / 10 cm	For the maximal intermalleolar distance
S = (89.3° - A) / 8.5°	For the cervical spine rotation (mean right/left)
Always with the additional condition $0 \leq S \leq 10$[a]	

[a] S=score, A=assessment.

For cervical rotation, tragus-to-wall distance and lumbar flexion, take the mean of the left and right measurements, if both are available. Otherwise, the available measurement will be used.

With regard of the lumbar flexion (modified Schober), values greater than 9.0 cm will be flagged as invalid and treated as if they were missing values.

When calculating observed case score, if one or more clinical measures are missing at a visit, then BASMI total score will be treated as missing at that visit. No imputation will be performed.

When performing LOCF imputation, following rules will be applied:

- If up to two clinical measures are missing at a visit, the previous value(s) for the clinical measure(s) will be carried forward to replace the missing value(s), assuming the previous value(s) were from a post-baseline assessment. If there are no previous post-baseline value(s) to carry forward, then the previous visit's total BASMI score will be carried forward, assuming the previous visit was a post-baseline assessment. Otherwise the LOCF score is missing for that visit.
- If more than two clinical measures are missing at a visit, then the previous visit's total BASMI score will be carried forward, assuming the previous visit was a post-baseline assessment. Otherwise the LOCF score is missing for that visit.

3.2.6 Total and nocturnal spinal pain NRS

Total and nocturnal spinal pain scores are based on following NRS:

- Pain in the spine due to AS (ie, "How much pain of your spine due to spondylitis do you have?")
- Pain in the spine at night due to AS (ie, "How much pain of your spine due to spondylitis do you have at night?").

The scores range from 0 to10, with lower scores indicating less pain.

3.2.7 Ankylosing Spondylitis Disease Activity Score (ASDAS)

The ASDAS is comprised of a number of assessments which are scored by the subject and physician and multiplied by a proven formula as listed:

0.121 × Back pain (BASDAI Q2 result, see Section 3.2.1)

0.058 × Duration of morning stiffness (BASDAI Q6 result)

0.110 × PtGADA (see Section 3.2.10)

0.073 × Peripheral pain/swelling (BASDAI Q3 result)

0.579 × (natural logarithm of the (CRP [mg/L] + 1))

Back pain, PtGADA, duration of morning stiffness and peripheral pain/swelling are all assessed on an NRS (0 to 10 units). If the CRP value is below the LLOQ, then it will be imputed as the midpoint between 0 and the LLOQ.

For Part A Baseline derivation of the ASDAS score the latest value prior to treatment will be used for each component. It will be not required that all components are measured at the same visit, e.g. PtGADA is available only at screening while the other components are available at Baseline then the ASDAS score will be derived combining the screening value of PtGADA and the baseline value of the other components.

For Part B Baseline derivation of the ASDAS score the latest value prior to randomization will be used for each component. It will be not required that all components are measured at the same visit.

For Flare Baseline derivation of the ASDAS score the latest value prior to first injection in Escape Part will be used for each component. It will be not required that all components are measured at the same visit.

When calculating observed case score, if one or more components are missing at a visit, then ASDAS will be treated as missing at that visit. No imputation will be performed.

When performing LOCF imputation, following rules will be applied:

- If only one component is missing at a visit, the previous value for that one component only will be carried forward to replace the missing value, assuming the previous value was a post-baseline assessment, and will be used to calculate the ASDAS score. If there is no previous post-baseline value available for that one component to carry forward, the previous visit's total ASDAS score will be carried forward, assuming the previous visit was a post-baseline assessment. Otherwise the LOCF score is missing at that visit.
- If more than one component is missing at a visit, the previous visit's total ASDAS score will be carried forward, assuming the previous visit was a post-baseline assessment. Otherwise the LOCF score is missing at that visit.

Disease activity is measured by binary response variables derived from ASDAS as follows:

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

ASDAS improvement is measured by binary response variables defined as follows:

- ASDAS-CII: ASDAS reduction (improvement) of ≥1.1 relative to Baseline
- ASDAS-MI: ASDAS reduction (improvement) of ≥2.0 relative to Baseline. As a sensitivity analysis, ASDAS-MI analyses will be repeated using a Modified ASDAS-MI: ASDAS reduction (improvement) of >=2.0 relative to Baseline or has the lowest score possible post-baseline (i.e. when CRP<LLOQ and all other components are 0, then the minimum ASDAS score is 0.636 to 3 decimal places).

3.2.8 Ankylosing Spondylitis Quality of Life (ASQoL)

The ASQoL, a validated disease-specific 18-item questionnaire, has been recently developed specifically for measuring health-related quality of life (HRQoL) in subjects with AS. The ASQoL has been used and has shown to be responsive in axSpA. The ASQoL score ranges from 0 to 18 with higher score indicating worse HRQoL.

If three or less items responses are missing, the ASQoL score will be imputed by dividing the sum score with the number of available responses and multiplying the result with 18. If more than three items responses are missing, the ASQoL score will be left missing.

3.2.9 Work Productivity Survey (WPS)

[redacted]

[redacted]

3.2.10 Patient's Global Assessment of Disease Activity (PtGADA, NRS)

Subjects will score their global assessment of their disease activity in response to the question "How active was your spondylitis on average during the last week?" using a NRS where 0 is "not active" and 10 is "very active".

3.2.11 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.

For the calculation of the SF-36 domain scores and the component summaries PCS and MCS, the scoring software Optum's PRO CoRE will be used. The norm-based scores (based on the US 2009 general population) will be utilized for analysis.

A response definition for the PCS and MCS is based on the within-subject MCID of 2.5 points. A subject will be considered a PCS (MCS) responder if the subject has an increase of ≥2.5 points from Baseline.

3.2.12 Maastricht Ankylosing Spondylitis Enthesitis Score (MASES)

The MASES comprises 13 items (bilateral costochondral 1, costochondral 7, anterior superior iliac spine, posterior iliac spine, iliac crest and proximal insertion of the achilles tendon sites, and the fifth lumbar vertebral body spinous process) (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 items are available, MASES will be treated as missing.

3.2.13 Swollen and Tender Joint Counts (44)

Tender and swollen joint counts 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 the 44 joints from the above list and ranges from 0 to 44. Artificial, ankylosed, and missing joints are excluded from swelling and tenderness assessments.

Each joint is scored as follows:

Table 3.2: Tender and swollen joint scoring

Grade	Swelling response (44)	Tenderness response (44)
0	None	None
1	Swelling present	Tenderness present

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 by 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 count will be set to missing for that visit.

3.2.14 Physician's Global Assessment of Disease Activity (PhGADA)

The Investigator will assess the overall status of the subject with respect to the axSpA signs and symptoms and the functional capacity of the subject using a VAS 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."

This assessment by the Investigator should be made without any knowledge of the PtGADA.

3.2.15 Spinal mobility

In addition to the assessments performed for the BASMI, additional spinal mobility assessments include:

- Occiput to wall distance
- Chest expansion

3.2.16 Health status (EQ-5D-3L)

The EQ-5D-3L is comprised of a 5-item health status measures 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-3L 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).

This instrument is to be completed by the subject.

3.2.17 Resource utilization

The following resource utilization data will be collected through UCB standardized modules:

- Concomitant medical procedures
- Health care provider consultations not foreseen by the protocol
- In-patient hospitalizations and emergency room visits (including length of stay)

A medical resource will be allocated to Part A as determined by the (start) date of the event. Specifically, a medical resource will be allocated to Part A if (start) date is on or after the date of first study medication in Part A and prior to or on the date of last study medication in Part A + 70 days and prior to the date of first study medication in Part B, if any.

A medical resource will be allocated to Part B as determined by the (start) date of the event. Specifically, a medical resource will be allocated to Part B if (start) date is on or after the date of first study medication in Part B and prior to or on the date of last study medication in Part B + 70 days. Medical resource with a (start) date after date of first escape full-dose CZP injection will be excluded from any summary tables of medical resource in Part B.

A medical resource will be allocated to Part B post-flare as determined by the (start) date of the event. Specifically, a medical resource will be allocated to Part B post-flare if (start) date is on or after the date of first escape full-dose CZP injection in Part B and prior to or on the date of last escape full-dose CZP injection in Part B + 70 days.

The same rules for (partially) missing start dates as for AEs will be applied to the resource use data.

For the same subject only one hospitalization will be considered if "start date of the second hospitalization - end date of the first hospitalization ≤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 in-patient hospitalizations and emergency room visits are not available during Part A, Part B and Part B post-flare the respective number variable will be set to 0.

See Section 12.1.2 for specific counting rules for the resource utilization.

3.2.18 Extra-articular assessments

The evolution of associated nonmusculoskeletal features including IBD, psoriasis and uveitis (including their severity), and flare rate history will be assessed as described in the protocol in the schedule of study assessments Table 5.1. For subjects experiencing a flare in Part B, corresponding assessments per visit are described in the protocol in the schedule of study assessments Table 5.2.

Incidence of uveitis flares, IBD exacerbations and psoriasis exacerbations will be provided from the extra-articular assessments.

3.2.19 Fecal and serum calprotectin

The calgranulins, S100A8/S100A9 and S100A12, are calcium-binding proteins of the S100 family. They are released from activated monocytes and granulocytes at local sites of inflammation during the early phase of the immune response. Extracellularly, they exert important pro-inflammatory effects, thereby providing stimulation and amplification of the innate immune reaction. Calgranulins can be measured in serum and stool, and have been found to be very sensitive markers of innate immune activation. Moreover, the S100A8/S100A9 heterodimer, also called calprotectin, has been established for a long time as a fecal marker of disease activity in inflammatory bowel disease.

Fecal and serum calprotectin levels measured as numerical continuous variables will be assessed from the stool and serum samples collected during the study.

3.2.20 Inflammatory Bowel Disease Questionnaire (IBD-Q)

The IBD-Q is a validated quality of life questionnaire, which combines an assessment of symptoms as a result of IBD, the way the subject has been feeling in general, and how their mood has been over the previous 2 weeks.

The IBD-Q contains 32 questions covering 4 domains of subjects' lives (bowel symptoms, systemic symptoms, emotional function, and social function).

Subjects will be asked to answer questions relating to the impact of their CD on their HRQoL over the 2 weeks prior to their study visit using the IBD-Q. The IBD-Q must be completed by the subject prior to start any study-related activities.

The method as detailed by Guyatt et al, 1989 will be used in the calculation of the IBD-Q score.

The total IBD-Q global score will be derived as the sum of the responses (from 1 to 7) to all 32 questions on the IBD-Q and can therefore range from 32 to 224. The IBD-Q score for each of the four categories (bowel symptoms, systemic symptoms, emotional function and social function) will be defined as the sum of the responses of the subset of questions of the IBD-Q as specified below:

- Bowel Symptoms Questions 1, 5, 9, 13, 17, 20, 22, 24, 26, 29
- Systemic Symptoms Questions 2, 6, 10, 14, 18
- Emotional Function Questions 3, 7, 11, 15, 19, 21, 23, 25, 27, 30, 31, 32
- Social Function Questions 4, 8, 12, 16, 28.

An IBD-Q response is defined as a clinically meaningful improvement in the total IBD-Q global score and corresponds to an increase in the total IBD-Q global score ≥16 points from Baseline. IBD-Q remission is defined as a total IBD-Q global score ≥170 points.

The following approach will be applied in the case of partially completed questionnaires.

The IBD-Q score for the four categories (bowel symptoms, systemic symptoms, emotional function and social function) will be calculated if only, at maximum, one of the items responses within the category is missing. The IBD-Q category score will be imputed by dividing the sum score with the number of available items responses and multiplying the result by the number of completed items within the category. If two or more items responses are missing within a category, the corresponding IBD-Q category score will be left missing.

The total IBD-Q global score will be calculated if four or less of the items responses are missing, as long as maximum one item per sub score is missing. The total IBD-Q global score will be imputed by dividing the sum score with the number of available items responses and multiplying the result by 32. If five or more items responses are missing, the corresponding total IBD-Q global score will be left missing.

Specific counting rules for answers provided to IBD-Q items

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:

1. In case of out of range answer (ie, an answer that does not correspond to any possible response proposed in the questionnaire, e. g. "?", "I don't know" or any value superior or inferior to the ones specified in the response options): the answer will be scored as "missing". However, in case the subject selects one of the proposed responses but adds a comment (for example "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 selects one of the proposed responses but adds a value superior or inferior to the ones specified in the responses options

(for example “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").

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

a. 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 of the 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 (ie the lowest item score for the IBD-Q).

b. If a response is marked between two allowed responses (for example, 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 (lowest) score will be retained.

3.2.21 Sustained Remission

Sustained remission is achieved when a subject has an ASDAS <1.3 at Week 32 or Week 36 (if ASDAS <1.3 at Week 32, it must be <2.1 at Week 36; if ASDAS <2.1 at Week 32, it must be <1.3 at Week 36) and an ASDAS <1.3 at Week 48.

3.2.22 Flare

A subject is considered to have experienced a flare if a subject has an ASDAS ≥2.1 at 2 consecutive visits or an ASDAS >3.5 at any visit during Part B up until Week 96.

If a subject has an intermittent missing ASDAS score (ie, a missing value at 1 visit but non-missing values at the previous and subsequent visit), then the missing value will be ignored and the occurrence of a flare will be based on the 2 available values.

If the ASDAS score is missing at 2 or more consecutive visits, then the subject will be considered to have experienced a flare (even if ASDAS scores are available after the missed visits).

There is also the possibility to report one unscheduled ASDAS between two scheduled ASDAS assessments. The same rules as for the scheduled ASDAS apply for the unscheduled ASDAS.

The unscheduled ASDAS will be taken into account for flare calculation for two consecutive visits if there are more than 13 calendar days between two consecutive ASDAS calculated dates.

If the ASDAS score is missing at 2 scheduled visits and unscheduled ASDAS entry present in between and not larger than 3.5, then the subject will not be considered to have experienced a flare.

Occurrence of a flare will be based solely on the observed ASDAS values and will not take into account any imputation of ASDAS. Occurrence of a flare will be determined based on the actual

data entered into the eCRF for the individual ASDAS components, not based on data from the interactive voice or web response system (IXRS).

If a subject prematurely withdraws from Part B at any point for any reason, then the subject will be considered to have experienced a flare.

If a subject was incorrectly identified as flared at any point in Part B due to late data entry in IXRS and went to escape Part B, then the subject will be considered to have experienced a flare.

Sensitivity analyses will be conducted using different approaches for definition of flare. These approaches are described hereafter.

3.2.22.1 Observed case

Using this approach, a subject will be considered to have experienced a flare only if a subject has an ASDAS ≥2.1 at 2 consecutive visits or an ASDAS >3.5 at any visit during Part B up until Week 96. Only the subjects with at least one ASDAS score available at any post-baseline visit during Part B until Week 96 will be evaluable using this approach.

3.2.22.2 Alternative handling of premature withdrawal

In this approach, a different handling of subject experiencing premature withdrawal from Part B at any point will be considered. Only subjects with a primary reason for premature withdrawal recorded as "adverse event" or "lack of efficacy" will be considered as having a flare. The other subjects experiencing premature withdrawal from Part B will not be considered as having a flare (unless other flare criteria are met).

3.2.22.3 Mixed model with repeated measures (MMRM) imputation of ASDAS

In this approach, MMRM imputation of missing ASDAS values at any visit in Part B up until Week 92 because of subject premature withdrawal, of subject experiencing a flare or of missing assessment at a given visit will be performed. Consequently, using this approach, a subject will be considered as having experienced a flare only if a subject has an ASDAS ≥2.1 at 2 consecutive visits or an ASDAS >3.5 at any visit during Part B up until Week 92, taking into account both observed and imputed values.

The MMRM analysis is a special case of the mixed-effects regression model (Siddiqui et al, 2009) and can be generally written as:

$$Y = \mathbf{X}\beta + \mathbf{Z}\nu + \varepsilon$$

where Y is the continuous variable of which the missing value is estimated; **X** is the n×p matrix of fixed effects in the model; **Z** is the n×r matrix of variables related to the random component in the model; β is the p×1 vector of regression coefficients for the fixed effects; ν is the r×1 vector of random subject effects; and ε is the n×1 vector of the random residuals in the model.

The MMRM imputation of missing ASDAS values will be performed using an ANOVA/ANCOVA type of MMRM with randomized treatment arm (PBO, CZP 200mg Q2W, CZP 200mg Q4W), visit (categorical, ie Week 48, 50, 52...), geographic region (North America, Western Europe, Eastern Europe and Asia) and mNY classification (mNY positive, mNY negative) included in the model.

The multiple visits for each subject are the repeated measures as random effect within each subject.

The covariance structure of the within-subject errors will be assumed to be unstructured. If the covariance structure does not allow for a stable solution, a variance component structure or alternatively an autoregressive first order structure will be used.

The resulting estimates of the regression model will be used for replacing the missing values of a variable. If one of the regressor variables is missing, the adjusted least-squares mean at the given visit will be used for replacing the missing value.

3.2.22.4 Referenced based multiple imputation of ASDAS

Missing data for the ASDAS scores at any visit in Part B up until Week 92 because of subject premature withdrawal, of subject experiencing a flare or of missing assessment at a given visit will be handled at each visit via reference based multiple imputation as described below:

This procedure will use an imputation model developed based on data from the PBO group only (Mallinckrodt, 2013). The reference-based multiple imputation assumes that the statistical behavior of the CZP and PBO-treated subjects after discontinuing study medication becomes that of the PBO-treated subjects. Data collected after discontinuation of the double-blind study treatment for both the CZP and PBO groups will be considered missing. Multiple imputations are used to replace missing outcomes for CZP- and PBO-treated subjects who discontinued using multiple draws from the posterior predictive distribution estimated from the PBO arm.

For subjects with a non-monotone (intermittent) missing pattern, Markov chain Monte Carlo (MCMC) will be used to impute their missing intermittent ASDAS scores by randomized treatment, geographic region and mNY classification. This will be done only once for each subject in order to provide a dataset with monotone missing (seed will be 221).

Data are processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcome data at visits t=1,…. T, where t=1 is Week 48 and t=T is Week 92

Initialization. Set t=1 (Part B Baseline visit)

Iteration. Set *t*=*t*+1. Create a data set combining records from CZP- and PBO-treated subjects with columns for covariates (geographic region and mNY classification) and outcomes at visits 1 to t. Outcomes for all CZP-treated subjects are set to missing at visit t and set to observed or previously imputed values at visits 1 to t-1. Outcomes for PBO-treated subjects are set to observed at visit t or observed or previously imputed values at visits 1 to t-1.

Imputation. Run MCMC to impute missing values for visit *t* using previous outcomes for visits 1 to *t*-1, geographic region and mNY classification. Note that only PBO data will be used to estimate the imputation model since no outcome is available for drug-treated subjects at visit *t*.

Repeat steps 2a-2d, 20 times with different seed values to create 20 imputed complete data sets (seeds will be 51, 52, 53, 54, 55, 56, 57, 58, 59, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 221).

Analysis. For each completed data set, the flare status will be calculated using the complete datasets (with no missing data). Analysis as described in Section 8.1.1 will be performed in order to provide treatment odds ratios, 95% CIs and the p-values from the proc MIanalyze.

If there are not enough PBO subjects available at the visits to impute the missing ASDAS values at the visits in Part B so that the model converted, no data will be presented in the corresponding table.

3.2.23 Time to flare

The time to flare will be defined as follows:

- For subjects who meet the criteria for flare, the flare date will be defined as:
  - For subjects who first meet the criterion "ASDAS ≥2.1 at 2 consecutive visits", the date of the second ASDAS score will be considered the flare date.
  - For subjects who first meet the criterion "ASDAS >3.5 at any visit", the date of the corresponding ASDAS score will be considered the flare date.
  - The date of the ASDAS score is defined as the date the corresponding BASDAI/PtGADA assessments were performed and the CRP sample(s) was (were) collected. If the assessments were not performed on the same day, a conservative approach will be used, and the latest day will be taken as calculated ASDAS day.
- Subjects who discontinue the study without Withdrawal Visit will be treated as having experienced a flare at the next planned ASDAS visit. The final available ASDAS score date + 32 (4 Weeks + visit window) will be used as flare date. If Week 48 or Week 50 are last ASDAS visits, Week 48/Week 50 ASDAS date + 18. (2 weeks + visit window) will be used as flare date.
- Subjects who discontinue the study will be treated as having experienced a flare at the date of their withdrawal visit, if the last ASDAS visit date is not equal with the withdrawal date. If the withdrawal date is equal to the last ASDAS visit date, the same approach as for subject without Withdrawal Visit will be used for deriving the flare date.
- Subjects with missing ASDAS scores at 2 or more consecutive visits will be treated as having experienced a flare at the date of the second visit where the ASDAS score is missing, irrespective of whether the subject meets the flare criteria at a later time point.
  - If the date of the second missing ASDAS visit is available, the date of the visit will be used.
  - If the date of the second missing ASDAS visit is not available, the last ASDAS date + 60 (8 weeks +visit window) will be used as flare date. If Week 48 is the last ASDAS visits, Week 48 ASDAS date + 32. (4 weeks + visit window) will be used as flare date. If Week 50 is the last ASDAS visits, Week 50 ASDAS date + 46. (6 weeks + visit window) will be used as flare date.
- For subjects who complete the study without meeting the criteria for flare, the censored date will be the last possible date of the event, i.e. Week 96.

The time to flare will be calculated in days as the difference between the two dates (date of flare or censoring date – date of randomization into Part B+ 1).

Sensitivity analyses will be conducted using different approaches for definition of time to flare. These approaches are described hereafter.

Observed case

Using this approach, a subject will be considered to have experienced a flare only if a subject has an ASDAS ≥2.1 at 2 consecutive visits or an ASDAS >3.5 at any visit during Part B up until Week 92.

Consequently, unlike for the main approach, subjects who discontinue the study will be censored at the date of their Withdrawal Visit and subjects with missing ASDAS scores at 2 or more consecutive visits will be censored at the date of the second visit where the ASDAS score is missing. In case a subject has missing ASDAS scores at 2 or more consecutive visits and meets the flare criteria at a later time point, the subject will be considered to have experienced a flare and the flare date will be defined as described earlier in this section.

Only the subjects with at least one ASDAS score available at any post-baseline visit during Part B will be evaluable using this approach.

Alternative handling of premature withdrawal

In this approach, a different handling of subject experiencing premature withdrawal from Part B at any point will be considered. Only subjects with a primary reason for premature withdrawal recorded as “adverse event” or “lack of efficacy” will be considered as having a flare. The other subjects experiencing premature withdrawal from Part B will not be considered as having a flare (unless other flare criteria are met).

Consequently, subjects with a primary reason for premature withdrawal recorded as “adverse event” or “lack of efficacy” will be treated as having experienced a flare at the date of their Withdrawal Visit. The other subjects experiencing premature withdrawal from Part B will not be considered as having a flare (unless other flare criteria are met) and will be censored at the date of their Withdrawal Visit.

Overview of different approaches for flare definition:

Table 3.3: Flare Definition and Time to Flare

Reason for flare	Event/Censoring day	Non-responder imputation	Observed Case	Alternative handling of premature withdrawal
ASDAS >=2.1 at 2 consecutive visits	Day of 2nd ASDAS	Event	Event	Event
ASDAS >3.5	Day of ASDAS	Event	Event	Event
Subject with Missing ASDAS at two consecutive visits	Planned date for the 2nd ASDAS visit	Event	Censored	Event
Subject discontinued study during Part B due to AE or Lack of efficacy	Day of discontinuation not equal ASDAS visit	Event	Censored	Event
Subject discontinued study during Part B due to other reason	Day of discontinuation not equal ASDAS visit	Event	Censored	Censored
No Flare (completed Part B)	Week 96	Censored	Censored	Censored
Incorrectly 'flared' due to delay in data entry	2nd ASDAS visit or planned date for the 2nd ASDAS visit if missing	Event	Censored	Censored

3.2.24 Subject age

Age in years will be calculated in the Study Data Tabulation Model (SDTM) data using the methodology defined in the UCB Interpretation of the SDTM Implementation Guide. For subjects with partial dates of birth (DOB), age will be calculated using the imputed DOB as provided by data management (DM). The imputation of partial DOBs by DM will be based on UCB data management standards.

3.2.25 Subject race

All subjects enrolled at French sites will not provide race data. Even in the case such data was provided by the sites in error, it will be ignored for tables and only presented in listings.

3.2.26 Anti-CZP antibody status

Determination of ADAb will be done using a validated screening, confirmation, and titration ADAb bridging assay. The immunogenicity data will be analyzed according to dedicated Bioanalytical Analysis plans.

3.2.27 Modified New York criteria

Subjects will be defined as having radiographic axial spondyloarthritis (r-axSpA) or non-radiographic axial spondyloarthritis (nr-axSpA) in this study based solely on the radiologic criterion. For simplification, r-axSpA will be referred to AS in the SAP and statistical outputs. If

a subject has sacroiliitis grade ≥2 bilaterally or sacroiliitis grade 3 to 4 unilaterally on x-ray then the subject will be considered to have AS/be modified New York (mNY) positive, otherwise they will be considered to have nr-axSPA/be mNY negative.

The mNY classification will be determined at Screening. MRI will be assessed centrally and by 2 independent readers. If the results from the readers are different, then the result of the adjudicator will be used as the adjudicator's decision was considered final for the inclusion of the subject in the study. If the modified NY criteria data recorded in the 'Modified NY Criteria for Ankylosing Spondylitis' eCRF differs from the MRI result reported by central review, the result from central review will be used in summary table instead of the result reported in the eCRF.

3.2.28 Modified ASAS Classification Criteria

If the results for CRP, human leukocyte antigen B27 (HLA-B27) or MRI on the 'Modified ASAS Classification Criteria for AS' eCRF differ from central laboratory results and/or the MRI result from central review, respectively, the result of the central laboratory and/or central review of MRI will be used in the summary table for modified ASAS classification criteria.

3.2.29 Past, prior and concomitant medications

Medication taken before the date of first study drug administration in Part A (regardless of continuing or stopping the medication) will be considered as prior medication in Part A.

Prior medication in Part A, stopping before the first study drug administration in Part A will be considered as 'past medication'.

Prior medication in Part A not stopped before first study drug administration in Part A will be considered as concomitant medication during Part A. Medication for which the start date is between the first date (including the date) of study drug administration in Part A and the date of the last study drug administration in Part A +14 days (including the date) and prior to the date of first study drug administration in Part B if any, will also be considered as concomitant medication during Part A.

Medication taken before the date of first study drug administration in Part B (regardless of continuing or stopping the medication) will be considered as prior medication in Part B.

Prior medication in Part B not stopped before first study drug administration in Part B will be considered as concomitant medication during Part B. Medication for which the start date is between the first date (including the date) of study drug administration in Part B and the date of the last study drug administration in Part B + 14 days (including the date) will also be considered as concomitant medication during Part B.

In subjects experiencing a flare in Part B, a concomitant medication during Part B taken before the date of flare (regardless of continuing or stopping the medication) will be considered pre-flare concomitant medication in Part B. Concomitant pre-flare medication in Part B not stopped before the date of flare will be considered as post-flare concomitant medication during Part B. In subjects experiencing a flare in Part B, medication for which the start date is after or on the date of flare will also be considered as post-flare concomitant medication during Part B.

3.3 Analysis time points

The analysis time point of interest for Part A is Week 48. Additional time points of interest for Part A are the Screening Visit, Week 0, Week 2, Week 4, Week 12, Week 24, Week 32 and

Week 36. For subjects who withdraw early during Part A, data recorded at the Withdrawal Visit will be assigned to what would have been the next scheduled visit for the particular assessment that is being summarized, unless otherwise specified. These data will be included in the summaries of the assigned visit and the "Early Withdrawal" visit if applicable. If a subject withdraws early during Part A but does not have a Withdrawal Visit in Part A, safety and efficacy data recorded at either the date of the last scheduled or unscheduled visit during the treatment Period, whichever is later, will be included in summaries of "Early Withdrawal" visit. Only post Part A Baseline assessments will be included in the "Early Withdrawal" visit.

The analysis time point of interest for Part B is Week 96. Additional time points of interest for Part B are Week 48, Week 50, Week 52, Week 54, Week 56, Week 58, Week 60, Week 62, Week 64, Week 66, Week 68, Week 70, Week 72, Week 74, Week 76, Week 78, Week 80, Week 82, Week 84, Week 86, Week 88, Week 90, Week 92 and Week 94. For subjects who complete Week 96, safety and efficacy data recorded at a Week 96/withdrawal Visit will be included in summaries of the Week 96 visit and the "Last/Withdrawal" visit if applicable. For subjects who withdraw from Part B prior to Week 96, data recorded at the Withdrawal Visit will be assigned to what would have been the next scheduled visit for the particular assessment that is being summarized, unless otherwise specified. These data will be included in the summaries of the assigned visit and the "Last/Withdrawal Visit" if applicable. If a subject withdraws early during Part B but does not have a Week 96/withdrawal Visit, then safety and efficacy data recorded at either the date of the last scheduled or unscheduled visit during the treatment Period, whichever is later, will be included in summaries "Last/Withdrawal Visit". For subjects experiencing a flare, the last visit prior to escape medication (e.g. Escape Week 0) will be included in the summaries of the "Last/Withdrawal Visit". Only post Part B Baseline assessments will be included in the "Last/Withdrawal Visit".

The analysis time point of interest for subjects experiencing a flare in Part B is Escape Week 12. Additional time points of interest for this set of subjects are Escape Week 0, Escape Week 2, Escape Week 4, Escape Week 6, Escape Week 8, Escape Week 10, Escape Week 24, Escape Week 36 and the final assessment visit as laid down for Week 96. For subjects who complete the final assessment visit as laid down for Week 96, safety and efficacy data recorded at this visit will be included in summaries of the "Last Visit (Week 96 Termination)" visit and the "Last/Withdrawal Escape" visit if applicable. For subjects who withdraw prematurely from the Escape Part, data recorded at the Withdrawal Visit will be assigned to what would have been the next scheduled visit for the particular assessment that is being summarized, unless otherwise specified. These data will be included in the summaries of the assigned visit and the "Last/Withdrawal Escape" visit if applicable. If a subject withdraws early during the Escape Part but does not have a Withdrawal Visit, then safety and efficacy data recorded at either the date of the last scheduled or unscheduled visit during the Escape Part, whichever is later, will be included in summaries "Last/Withdrawal Escape" visit. Only post Flare baseline assessments will be included in the "Last/Withdrawal Escape" 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. In case of repeated assessments at a particular visit, the last assessment will be used for visit presentation in by-visit summaries (except unscheduled repeats of Week 96).

PK and ADAb

For ADAb data the same approach as for the efficacy analysis will be used, i.e for withdrawals to assign the withdrawal value to the next scheduled visit for the assessment. However, for PK analysis the withdrawal PK result will be only included in summaries of "Withdrawal" and the "Withdrawal Escape" and not assigned to the next scheduled visit. Furthermore, the Withdrawal summaries visit will not summarize any last visit assessments.

In addition, in the case that a sample is collected one or more days following the scheduled visit date, in which the drug was administered, the PK results for that sample will not be summarized as part of that visit.

3.3.1 Relative day

The relative day will be included in different listings as appropriate.

For Part A, relative day, identified as "relative day [a]", will be calculated relative to the first CZP administration in Part A. Relative day [a] will be calculated as follows:

Relative day [a] of date X = date X - Date of first CZP administration in Part A+1 if date X is on or after date of first CZP administration in Part A.

Relative day [a] of date X = date X - Date of first CZP administration if date X is before date of first CZP administration in Part A. Relative day [a] before first CZP administration in Part A will have the prefix "-".

Relative day [a] after last CZP administration in Part A will have the prefix "+", and will be calculated from the date of last CZP administration in Part A (date X - date of last CZP administration in Part A).

For Part B, relative day, identified as "relative day [b]", will be calculated relative to the first study medication administration in Part B. Relative day [b] will be calculated as follows:

Relative day [b] of date X = date X - Date of first study medication administration in Part B+1 if date X is on or after date of first study medication administration in Part B.

Relative day [b] of date X = date X - Date of first study medication administration if date X is before date of first study medication administration in Part B. Relative day [b] before first administration of study medication in Part B will have the prefix "-".

Relative day [b] after last study medication administration in Part B will have the prefix "+", and will be calculated from the date of last study medication administration in Part B (date X - date of last study medication administration in Part B).

For subjects experiencing a flare and starting escape therapy, an additional relative day, identified as "relative day [f]", will be calculated relative to the first study escape therapy administration in Part B. This additional relative day [f] will be calculated as follows:

Relative day [f] of date X = date X - Date of first study escape therapy administration in Part B+1 if date X is on or after date of first escape therapy administration in Part B.

Relative day [f] of date X = date X - Date of first escape therapy administration if date X is before date of first escape therapy administration in Part B. Relative day [f] before first escape therapy administration in Part B will have the prefix "-".

Relative day [f] after last escape therapy administration in Part B will have the prefix "+", and will be calculated from the date of last escape therapy administration in Part B (date X - date of last escape therapy administration in Part B).

Calculations of "Relative Day" will not include partial dates, but will be left blank in these instances.

3.3.2 End date of the Treatment Period

The end date of the treatment Period will be either Week 48 for subjects who completed Part A and did not meet the criteria to be randomized into Part B, Week 96 for subjects completing the treatment Period of Part B, or the date of the Withdrawal Visit for subjects who discontinued during the treatment Period (whether Part A or Part B). If a subject does not have a Week 96/Withdrawal Visit or has visits after Week 96 due to escape treatment, then either the date of the last scheduled or unscheduled visit during the treatment Period or the date of last known dose of study drug during the treatment Period, whichever is later, will define the end date of the treatment Period.

3.4 Definition of Baseline values

The study baseline value for a subject noted "Part A Baseline" will be defined as the latest measurement for that subject up to and including the day of administration of first study medication in Part A, unless otherwise stated. If a Part A Baseline measurement is missing or not collected, and a Screening value is available, the Screening value will be utilized as Part A Baseline instead. If no measurement is available prior to receiving study medication in Part A, the Part A Baseline value is treated as missing.

For subjects randomized into Part B, a "Part B Baseline" will be defined as the latest measurement for that subject up to and including the day of randomization, unless otherwise stated. If a Part B Baseline measurement is missing or not collected, and a post-Baseline Part A value is available, the post-Baseline Part A value will be utilized as Part B Baseline instead. If no post-Baseline Part A measurement is available prior to randomization, the Part B Baseline value is treated as missing.

For subjects experiencing a flare in Part B, a "Flare Baseline" will be defined as the latest measurement for that subject up to and including the day of administration of first escape treatment in Part B, unless otherwise stated. If a Flare Baseline measurement is missing or not collected, and a post-Baseline Part B value is available, the post-Baseline Part B value will be utilized as Flare Baseline instead. If no post-Baseline Part B measurement is available prior to receiving study medication in Escape Part, the Flare Baseline value is treated as missing. If a subject does not receive study medication in Escape Part, the last available post-Baseline Part B measurement will be used as Flare Baseline.

For each baseline definition, if there is evidence that measurements taken on the same day as administration of first study medication (in Part A/Part B) or first escape treatment (for subjects experiencing a flare in Part B) were actually taken after this administration, then only values strictly prior to that date (including screening assessment visits if appropriate) will be used for that subject. When the date of first dose is derived, it should be based on the first injection of study treatment, regardless of whether it is an active treatment.

Cases where the baseline value cannot be derived will be reviewed at the data evaluation meeting prior to database lock to determine an appropriate imputed value to use as substitute.

3.5 Protocol deviations

Important protocol deviations will be predefined in the AS0005 Protocol Deviation Specification (PDS) document. Pre-analysis data evaluation meetings will be held prior to the Part A and Part B database locks.

The aim of the pre-analysis data reviews are to review the Data Management snapshot and draft and accept the list of protocol deviations which will then be used to determine exclusion from analysis sets as appropriate. These protocol deviations will be assessed as important by a panel including the clinical project manager, lead clinical development representative, study statistician, and other appropriate clinical study team members using the criteria pre-specified within the PDS document.

3.6 Analysis sets

There will be 11 analysis sets used within this study.

3.6.1 Enrolled Set

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

3.6.2 Open-Label Set

The Open-Label Set (OLS) will consist of all subjects who receive at least 1 dose of study medication in the Open-Label Period of the study (Part A).

3.6.3 Randomized Set

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

3.6.4 Safety Set

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

3.6.5 Safety Set Part B

The Safety Set Part B (SSB) will consist of all subjects in the RS who have received at least 1 dose of study medication in the Double-Blind Period of the study (Part B).

3.6.6 Full Analysis Set

The Full Analysis Set (FAS) will be used to evaluate the sensitivity of the results of the primary efficacy analysis. The FAS will consist of all subjects in the RS who have received at least 1 dose of study medication in Part B, and have valid ASDAS measurements at Week 48 and at least 1 time point following Week 48.

3.6.7 Per Protocol Set

The Per Protocol Set (PPS) will be used to evaluate the sensitivity of the results of the primary efficacy analysis. The 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.

3.6.8 Pharmacokinetic Set B

The Pharmacokinetic Set B (PKSB) will consist of all subjects from the SSB who provide at least 1 PK sample during Part B.

3.6.9 Flared Set

The Flared Set (FS) will consist of all subjects from the RS experiencing a flare in Part B. Since IXRS will be used to determine whether flare criteria (ASDAS>3.5 once or >2.1 on 2 consecutive visits, or ASDAS missing on 2 consecutive visits) are met during the course of the study, the population will be based on the IXRS data and whether they enter the Escape Part according to IXRS.

3.6.10 Escape Therapy Set

The Escape Therapy Set (ETS) will consist of all subjects from the FS who have received at least 1 dose of escape therapy.

3.6.11 Part B Full-dose Set

The Part B Full-Dose Set (FDS) will include all subjects from the RS who ever received a dose of CZP Q2W during Part B (including escape treatment).

3.6.12 General use of the Analysis Sets

Efficacy summaries for variables collected in Part A will be based on the OLS. All efficacy analyses for Part B will be performed using the RS. The FAS and PPS will be used for a sensitivity analysis on the primary endpoint only. The other efficacy variables for subjects experiencing a flare in Part B will be based on the FS.

All safety summaries in Part A will be based on the SS. All safety summaries in Part B will be based on the SSB. The other safety variables for subjects experiencing a flare in Part B will be based on the ETS.

Pharmacokinetic and immunogenicity summaries will be based on PKSB.

3.7 Treatment assignment and treatment groups

During Part A, all subjects will receive the same treatment consisting of 3 loading doses of CZP subcutaneous 400mg Q2W, followed by CZP 200mg Q2W. In the tabular summaries for Part A, all subjects will be summarized together, under a ”All Subject” column. Summary tables may present subjects overall as well as broken out by axSpA subpopulation nr-axSpA and AS.

During Part B, subjects are expected to receive their treatment as randomized (PBO, CZP 200mg Q2W, CZP 200mg Q4W), unless escape treatment is required as a result of a flare. In the tabular summaries for Part B, subjects will be summarized according to their randomized or actual treatment group. Select summary tables will also present subjects broken out by axSpA subpopulation. In the tabular summaries for subjects experiencing a flare during Part B, subjects will be summarized according to their randomized treatment group; an overall column may also be included in select summary tables of safety data for these subjects.

Treatment assignments for the RS, FAS, PPS and FS will be according to randomization. Treatment assignments for the SSB, PKSB and ETS will be according to the actual treatment subjects received in Part B.

3.8 Center pooling strategy

Due to the small number of subjects expected at each center, centers will be pooled into regions (North America, Western Europe, Eastern Europe and Asia) for analysis purposes. The list of countries assigned to each region is provided in the table below.

Table 3.4: List of countries in each region

Region	List of countries
North America	USA
Western Europe	Belgium France Germany Netherlands Spain United Kingdom
Eastern Europe	Bulgaria Czech Republic Hungary Poland Romania
Asia	Turkey Taiwan

3.9 Coding dictionaries

Medical history and AEs will be recorded and coded using version 19.0 of the Medical Dictionary for Regulatory Activities (MedDRA®) criteria. Medications will be coded using the version from September 2015 of the World Health Organization Drug Dictionary (WHO-DD). Medical procedures will not be coded.

3.10 Changes to protocol-defined analyses

The SAP specifies which Baseline value is applicable for each analysis. The Flare Baseline is defined as the latest measurement for that subject up to and including the day of administration of first escape treatment in Part B instead of the visit in which the flare occurred.

In addition to the protocol, a modified version for ASDAS-MI is added for sensitivity analysis of ASDAS-MI.

Furthermore, analysis for IBD-Q response and IBD-Q remission using Part A Baseline and Part B Baseline as a reference and SF-36 PCS response and SF-36 MCS response using Part A Baseline as a reference are added.

Exploratory statistical comparisons for WPS will be based on MMRM instead of nonparametric bootstrap-t method.

ASDAS-Moderate Disease changed to ASDAS-Low Disease activity (Machado et al. 2018). LOCF analysis are removed from Part B efficacy tables.

MMRM will be used instead of ANCOVA based on LOCF for

- Total spinal pain
- Nocturnal spinal pain
- PtGADA
- Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
- SF-36 PCS, MCS, and physical function domain
- Fatigue NRS
- ASQoL.

The primary purpose of the study is to evaluate treatment options of axSpA patients after being in sustained remission. Hence, one of the objectives is to evaluate the PK and immunogenicity of these patients. The CZP plasma concentration and ADAb titer of the patients that did not reach sustained remission are therefore not analysed, the Pharmacokinetic Set A (PKSA) as described in the protocol has been removed from the SAP.

4 STATISTICAL/ANALYTICAL ISSUES

4.1 Adjustments for covariates

The randomization stratification factors of geographical region and the mNY criteria as defined in Section 3.2.27 will be used as covariates in the specified statistical models. The original values of the stratification factors used at the time of randomization will be recorded into IXRS. Of note in case any errors to the stratification factors is detected, no updates will be done to the IXRS data. However, the correct values of the stratification factors (including any corrections post randomization) will be used for the analyses.

4.2 Handling of dropouts or missing data

Post-flare efficacy data (except radiographic data) for subjects experiencing a flare (in any treatment group) will be treated as missing in Part B analyses. The pattern of missingness for these variables is assumed to be missing at random (MAR). For the summaries of data collected at post-flare visits in subjects experiencing a flare, there will be no adjustment for missing values and only observed case descriptive summaries will be provided.

For the primary efficacy variable (percentage of subjects in Part B who do not experience a flare), an approach based on the principles of NRI will be used to impute missing values. Specifically, if a subject withdraws from Part B at any point, they will be considered as having not responded to treatment. Therefore, they will be assumed to have experienced a flare during Part B for analysis purposes. Sensitivity analyses will be performed to assess the robustness of the results obtained using the NRI approach. Refer to Section 3.2.22 of this SAP for further details.

For the time to flare endpoint, the rules as described above for the primary endpoint will be applied to account for missing data when determining whether or not a subject experienced a

flare. Similarly, sensitivity analyses will also be performed to assess the robustness of the results obtained using the NRI approach. Refer to Section 3.2.23 of this SAP for further details.

For all analyses of binary endpoints (sustained remission, ASDAS-MI, ASDAS-CII, modified ASDAS-MI, ASAS20, ASAS40, ASAS5/6, ASAS PR, BASDAI50 response, IBD-Q remission, IBD-Q response, SF-36 PCS response, SF-36 MCS response) NRI will be used. Thus, unless mentioned otherwise (see Section 3.2.2 for ASAS20, ASAS40, ASAS5/6 and ASAS PR), a subject having missing data for the time point assessed will be conservatively counted as a non-responder. This will be done whether the data is missing, the subject discontinued prior to the time point assessed, or the data is considered missing due to start of escape treatment. For the following continuous endpoints in Part A, imputation of missing data will be performed using the LOCF approach:

- ASDAS
- BASDAI
- Fatigue item of BASDAI
- BASFI
- BASMI
- Nocturnal spinal pain score
- PtGADA
- CRP

In addition, ASDAS disease activity levels (ASDAS-ID, ASDAS-LD, ASDAS-HD and ASDAS-vHD) will be imputed based on the imputed values of ASDAS.

The LOCF imputation will be done whether the data is missing, the subject discontinued prior to the time point assessed, or the data is considered missing due to start of escape treatment.

Unless otherwise stated, the following guidelines apply for the LOCF analyses:

- For missing post-baseline Part A assessments, only carry forward earlier Part A post-Baseline values. Screening and Part A Baseline values do not get carried forward to the Part A post-Baseline visits.
- For missing post Week 48 Part B assessments,
  - if the subject did not experience a flare in Part B, only carry forward earlier Part B post-Week 48 values,
  - if the subject experienced a flare in Part B, the assessment at Escape week 0 will be treated as the last assessment in Part B prior to escape treatment. For any missing post 48 Part B assessments up to the week of flare (the later visit if flare was based on 2 visits), carry forward earlier Part B post-Week 48 values. Assessments whilst on escape treatment are not to be used, therefore the assessment at Escape week 0 will be carried forward for all subsequent visits until Week 96. E.g if the subject experienced a flare at week 72 then the Escape Week 0 assessment will be carried forward for the assessments after week 72 in Part B (76, 80, 84, 88, 92 and 96)

  - Week 48 values do not get carried forward to the Part B post-Baseline visits.

- Assessments during escape treatment will not be carried forward.
- The following continuous endpoints in Part B will be compared between treatment groups using a mixed model for repeated measures (MMRM). The pattern of missingness for these variables is assumed to be MAR. Subjects who enter Part B will have already demonstrated sustained remission through 48 weeks on CZP 200mg Q2W. Many of these subjects will be randomized to receive placebo and will, therefore, be more likely to experience a flare. Post-flare data for subjects experiencing a flare (in any treatment group) will be treated as missing for the MMRM analysis. It is further assumed that most subjects who discontinue at this late stage in the study (i.e. during Part B) will do so as a result of reduced efficacy. Therefore, missing efficacy data due to either flare or study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. For the following continuous endpoints in Part B, imputation of missing data will be performed using the MMRM approach:
- ASDAS
- BASDAI
- Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
- Fatigue item of BASDAI
- BASFI
- BASMI
- Total spinal pain score
- Nocturnal spinal pain score
- ASQoL
- PtGADA
- SF-36 physical functioning score
- SF-36 PCS
- SF-36 MCS
- WPS

The below rules will be applied for imputation of partial and missing start date of medications:

- If start day is missing, but month and year are present, start date will be replaced by the first day of the month of that year or by the day of start of study medication in Part A, if study medication in Part A was first administered in that month and year (exception: if in this case start date is after stop date, also the first day of the months will be utilized).
- If start day and month are missing and year is present, start date will be replaced by the first of January of that year or by the day and month of start of study medication in Part A, if study medication in Part A was first administered in that year (exception: if in this case start date is after stop date, also the first day of January will be utilized).

- If start day, month, and year are missing, start date will be replaced by the date of first study drug administration in Part A or by the first day of January of the year of the end date, if the first administration date is after the end date.

If medications are not ongoing (stopped) and the stop date is incomplete or missing, the following rules will be applied:

- If stop day is missing, but month and year are present, stop date will be replaced by the last day of the months or by the day of the last study drug administration in Part B if the last study drug administration in Part B occurs in that months and subject did not experience a flare. In subjects experiencing a flare in Part B and start escape therapy, stop date will be replaced by the last day of the months or by the day of the last administration of escape therapy if the last administration of escape therapy occurs in that month. For subjects not randomized into Part B, stop date will be replaced by the last day of the months or by the day of the last study drug administration in Part A if the last study drug administration in Part A occurs in that month. If in result of imputation, the stop date is before start date, also the last day of the months will be utilized.
- If stop day and months are missing and year is present, stop date will be replaced by last day of December or by the day and months of the last study drug administration in Part B if the last study drug administration in Part B was in that year and subject did not experience a flare. In subjects experiencing a flare in Part B and start escape therapy, stop date will be replaced by last day of December or by the day and months of the last administration of escape therapy if the last administration of escape therapy occurs in that year. For subjects not randomized into Part B, replacement by the day and months of the last study drug administration in Part A will be performed if the last study drug administration in Part A was in that year. If in result of imputation, the stop date is before start date, also the last day of December will be utilized.
- If stop day, months, and year are missing, stop date will be replaced by the date of the last visit in the double-blind period or by start date, if the last visit in the double-blind period date is before the start date.

The below rules will be applied for imputation of partial and missing AE start dates:

- If start day is missing, but month and year are present, start date will be replaced by the first day of the month of that year or by the day of start of study medication in Part A, if study medication in Part A was first administered in that month and year (exception: if in this case start date is after stop date, also the first day of the month will be utilized).
- If start day and month are missing and year is present, start date will be replaced by the first of January of that year or by the day and month of start of study medication in Part A, if study medication in Part A was first administered in that year (exception: if in this case start date is after stop date, also the first day of January will be utilized).
- If start day, month, and year are missing, start date will be replaced by the date of first study drug administration in Part A or by the first day of January of the year of the end date, if the first administration date is after the end date.

The below rules will be applied for imputation of partial and missing medical history start dates:

- If start day is missing, but month and year are present, start date will be replaced by the first day of the month of that year or by the day of start of study medication in Part A, if study medication in Part A was first administered in that month and year (exception: if in this case start date is after stop date, also the first day of the month will be utilized).
- If start day and month is missing for one item diagnosis (medical history of any of the preferred terms ‘Axial Spondyloarthritis’, 'Spondyloarthropathy', 'Ankylosing spondylitis' , or 'Spondylitis') or symptom (medical history preferred term of ‘Back pain’, ‘Inflammatory pain’, ‘Spinal Pain’, ‘Lumbar spine pain’) and the year is similar, the missing month will be imputed by the month of the other item.
- If only one item diagnosis or symptom is present and if start day and month are missing and year is present, start date will be replaced by the first of January of that year or by the day and month of start of study medication in Part A, if study medication in Part A was first administered in that year (exception: if in this case start date is after stop date, also the first day of January will be utilized).
- If start day, month, and year are missing, start date will be replaced by the date of first study drug administration in Part A or by the first day of January of the year of the end date, if the first administration date is after the end date.

4.3 Interim analyses and data monitoring

Main focus of this interim analysis will be the efficacy and safety up to Week 48. An interim freeze of the database will be performed once all subjects have completed Week 48 of Part A. Part A data will be reported in corresponding Part A tables and listings. All blinded data will remain blinded. In the disposition tables, treatment will remain blinded and be presented according to a dummy treatment assignment for Part B. For the listing only open-label data from Part A will be included. The purpose of the interim analysis is to check whether the assumptions on efficacy of CZP made for the protocol were correct, and to potentially present the results at scientific meetings. No CSR will be written. The following assessments will not be included in the Interim Analysis:

- MRI Efficacy scores (SPARCC (Sacroiliac joint), Spine ASspiMRI-a in Berlin Modification score)
- Work Productivity Questionnaire
- EQ-5D-3L
- Fecal and serum calprotectin
- Correlation between calprotectin and IBD, CRP
- SF36 (domains, PCS, MCS)
- Resource utilization
- PK
- ADAb

The scope of the Interim Analysis will not be covered in a separate Analysis Plan.

No specific data monitoring, steering, or evaluation committee is planned for this study.

The following variable will be classed as unblinding during the full study. Therefore, these variables will be kept blinded until study unblinding after database lock:

- Randomized treatment group
- CRP (Post Week 48)
- MRI Efficacy scores (Post Week 48)
- PK and ADAb (Post Week 48).

4.4 Multicenter studies

Individual center results will not be directly presented, however as discussed in Section 4.1 geographical region is used as a covariate and some outputs will be presented by geographical region. Centers will be grouped in the following regions: North America, Western Europe, Eastern Europe and Asia.

4.5 Multiple comparisons/multiplicity

The statistical analysis of the primary efficacy variable will account for the testing of multiple doses by using a fixed sequence testing procedure. The proportion of subjects who do not experience a flare will be compared between each CZP dose and PBO. The predefined order of hypotheses testing, each at a 2-sided 0.05 significance level for the comparison between the given CZP dose and PBO, will be performed in the sequence shown below:

1. CZP 200mg Q2W vs PBO
2. CZP 200mg Q4W vs PBO

The second test will be performed irrespective of whether the first test is significant at the 0.05 level or not. However, the second test will be interpreted as statistically significant only if the first test is significant at the 0.05 level as well.

No adjustment for multiplicity will be considered when analysing the secondary efficacy variables. The p-values calculated in conjunction with those variables will be considered as nominal.

4.6 Use of an efficacy subset of subjects

The primary efficacy analysis from Part B will be repeated using the FAS and the PPS as a supportive analysis.

4.7 Examination of subgroups

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

- mNY classification (mNY positive, mNY negative)
- Geographical region (North America, Western Europe, Eastern Europe and Asia)
- Age (<30 and ≥30 years)
- Gender

- Race (white and non-white)
- Baseline CRP level (≤upper limit of normal (ULN), >ULN)
- ADAb status (Negative, Positive <512, Positive 512-<1024, Positive >=1024).

These analyses will be based on descriptive statistics only.

5 STUDY POPULATION CHARACTERISTICS

5.1 Subject disposition

The number and percentage of subjects who entered the study, completed Part A, discontinued from Part A with reasons for discontinuation, were randomized into Part B, completed Part B (without flare, after flare) and discontinued Part B as well as started Escape Part, completed Escape Part and discontinued Escape Part with reasons for discontinuation will be presented.

Subjects who completed Week 48 without SFU are considered as completed Part A. Subjects completed Week 96 in Part B or in Escape Part without SFU visit are considered as completed Part B or completed Escape Part, respectively.

Subjects are considered to have started Escape Part, if they experienced a flare and have been moved to escape treatment via IXRS system (even if they do not receive any escape treatment).

This summary will be based on the ES and will present subjects overall and broken out by axSpA subpopulation.

Discontinuations due to AEs will also be presented. This summary will be based on the ES and will present subjects overall and broken out by axSpA subpopulation.

The disposition of subjects screened by geographical region and overall, including the number of subjects in each analysis set will be summarized. This summary will be based on the ES and will present subjects overall (ie not broken out by axSpA subpopulation).

A summary table of reasons for screen failures will be provided.

The following listings will be produced based on the ES:

- A listing of subject analysis sets
- A listing of subjects who did not meet study eligibility criteria
- A listing of subject disposition

The following listings will be produced based on the OLS:

- A listing of study discontinuation
- A listing of visit dates
- A listing of important protocol deviations (including any reasons for exclusion from the PPS)

5.2 Protocol deviations

Important protocol deviations that are likely to have an impact on the subject's rights, safety, well-being, and/or on the validity of the data for analysis will be identified during the blinded data evaluation meeting.

The identified important protocol deviations will be classified by pre-specified items (e.g., informed consent, inclusion and exclusion criteria, IP administration/study treatment, withdrawal criteria, disallowed medications, procedures/tests, visit schedule, safety and other) as well as study period of occurrence (Part A, Part B) and summarized in Part A using the OLS, in Part B using the RS and overall using the OLS.

All important protocol deviations will be provided in a data listing using the OLS.

6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

6.1 Demographics

The following demographic variables will be summarized by axSpA subpopulation and for all subjects combined using the OLS. This presentation will also be produced for the RS by treatment group.

Continuous variables: age (years), height (cm), weight (kg) and body mass index (BMI; kg/m2)

Categorical variables: age (<30 years, ≥30 years and also ≤18, 19 to 24, 25 to 34, 35 to 44, ≥45), gender (male, female), racial group (American Indian/Alaskan Native, Asian, Black, Native Hawaiian or Other Pacific Islander, White, Other/Mixed), ethnicity (Hispanic or Latino, Not Hispanic or Latino), BMI category (<18.5, 18.5 - <25, 25 - <30, 30 - <35, 35 - <40, ≥40) and region and country as per Table 3.4 .

All demographic details will be listed using the OLS.

6.2 Other Baseline characteristics

Lifestyle (alcohol, caffeinated beverages) and Tobacco usage at Baseline will be summarized by axSpA subpopulation and for all subjects combined using the OLS.

The responses to the 4 criteria for Modified NY Criteria for AS will be summarized by axSpA subpopulation and for all subjects combined in frequency tables using the OLS. This table will be also produced for the RS by treatment group.

Similarly, the responses to the following modified ASAS classification criteria for AS will be summarized by axSpA subpopulation and for all subjects combined in frequency tables using the OLS. This table will be also produced for the RS by treatment group.

- back pain of ≥3 months and subject's age at onset <45 years
- inflammatory back pain
- arthritis
- enthesitis
- uveitis
- dactylitis
- psoriasis
- Crohn´s disease (CD)/ulcerative colitis

- HLA-B27
- elevated CRP
- presence of sacroiliitis on imaging including imaging method

History of extra-articular manifestations will be summarized by axSpA subpopulation and overall in frequency tables using the OLS. This table will be also produced for the RS by treatment group.

- History of uveitis
- History of IBD
- History of psoriasis

Frequency tables by axSpA subpopulation and overall using the OLS will also be provided for:

- History of latent tuberculosis
- Tuberculosis test
- Chest x-ray
- Sacroiliac joint x-ray

All baseline characteristics described above as well as baseline extra-articular manifestations will be listed using the OLS.

For subjects randomized to Part B, frequency tables by treatment group using the RS and overall will be provided for the randomization stratification factors:

- Geographic region
- mNY classification

Summaries of the randomization stratification factors will include both the original values of the stratification factors (ie, as used for randomization) as recorded in the IXRS and the final values of the stratification factors (ie, updated values in case any modifications were made after randomization). The randomization stratification factors will also be listed using the RS.

6.3 Medical history and concomitant diseases

Previous and ongoing medical history will be summarized by MedDRA® system organ class (SOC) and preferred term (PT).

Previous and ongoing medical history will be summarized by axSpA subpopulation and for all subjects combined in Part A using the SS, by treatment group and for all subjects combined in Part B using the SSB, and will be listed using the SS.

Time since diagnosis of disease and symptom duration will be summarized by axSpA subpopulation and for all subjects combined in Part A using the OLS, by treatment group and for all subjects combined in Part B using the RS. Time since diagnosis will also be categorized into ≤2 and >2 years, and symptom duration will be categorized into ≤3 and >3 years.

Time since diagnosis of disease will be defined as: Date of first study medication administration minus earliest start date of the medical history of any of the preferred terms ‘Axial Spondyloarthritis’, 'Spondyloarthropathy', 'Ankylosing spondylitis', or 'Spondylitis' . If a subject

does not have a history of any of the above 4 preferred terms, then the time since diagnosis of disease will be set equal to the symptom duration.

Symptom duration will be calculated as:

(Date of first study medication administration – Start date of symptoms)/365.25

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 for concomitant medications in Section 4.2):

- Subjects with a medical history preferred term of 'Back pain', 'Inflammatory pain', 'Spinal Pain', 'Lumbar spine pain' and a start date present will use the earliest start date of these symptoms as the start date of symptoms.
- Otherwise, subjects with a preferred term of 'Axial spondyloarthritis', 'Spondyloarthropathy', 'Ankylosing spondylitis', or 'Spondylitis' 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.
- If the resulting symptom duration is unknown or is below 3 months, the imputed start date will be set to 3 months and 1 day prior to the start of treatment so that the symptom duration is imputed as 3 months.

Time since diagnosis and symptom duration as described above will be listed using the OLS.

6.4 Past, prior and concomitant medications

Past medication summaries will be generated for disease-modifying antirheumatic drugs (DMARDs), non-steroidal anti-inflammatory drugs (NSAIDs), and anti-tumor necrosis factors (anti-TNFs) by anatomical therapeutic chemical (ATC) code (level 3 (4-digit) decode, level 4 (5-digit) decode, and PT). See Section 3.2.29 for definition of past, prior and concomitant medications.

For Part A, prior DMARDs and NSAIDs will be summarized by WHO-DD ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and PT). 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-TNFs. Summaries will be presented by axSpA subpopulation and overall using the SS.

For Part A, prior medications (except DMARDs and NSAIDs) will be summarized by WHO-DD ATC code (level 2 (3-digit) and level 3 (4-digit) decode. Summary will be presented by axSpA subpopulation and overall using the SS.

For Part A, concomitant DMARDs and NSAIDs will be summarized by WHO-DD ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and PT). Summaries will be presented by axSpA subpopulation and overall using the SS.

For Part A, concomitant medication (except DMARDs and NSAIDs) will be summarized by WHO-DD ATC code (level 2 (3-digit) and level 3 (4-digit) decode. Summary will be presented by axSpA subpopulation and overall using the SS.

For Part B, prior DMARDs and NSAIDs will be summarized by WHO-DD ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and PT). Summaries will be presented by treatment group and overall using the SSB.

For Part B, prior medications (except DMARDs and NSAIDs) will be summarized by WHO-DD ATC code (level 2 (3-digit) and level 3 (4-digit) decode. Summary will be presented by treatment group and overall using the SSB.

For Part B, concomitant DMARDs and NSAIDs will be summarized by WHO-DD ATC code (level 3 (4-digit) decode, level 4 (5-digit) decode, and PT). Summaries will be presented by treatment group and overall using the SSB. Additionally, in subjects experiencing a flare in Part B, separate summaries by treatment group using the ETS will be provided presenting pre-flare and post-flare concomitant medications.

For Part B, concomitant medications (except DMARDs and NSAIDs) will be summarized by WHO-DD ATC code (level 2 (3-digit) and level 3 (4-digit) decode. Summary will be presented by treatment group and overall using the SSB. Additionally, in subjects experiencing a flare in Part B, separate summaries by treatment group using the ETS will be provided presenting pre-flare and post-flare concomitant medications.

Prohibited medication use and date of first usage will be determined and documented during the pre-analysis review meeting prior to the database lock.

The number and percentage of subjects who used prohibited medication will be summarized by axSpA subpopulation and overall in Part A using the SS and by treatment group and overall in Part B using the SSB. Separate summaries in subjects experiencing a flare in Part B by treatment group using the FS will be provided presenting pre-flare and post-flare prohibited medication. An overall summary of subjects who used prohibited medication in Part A or Part B will be provided using the SS.

Prohibited medications will be assigned to Part A or Part B (overall, pre-flare, post-flare) using the same rules as outlined for concomitant medications in Section 3.2.29 .

Prior and concomitant medications and prohibited medications will be listed using the SS.

7 MEASUREMENTS OF TREATMENT COMPLIANCE

Treatment compliance will be calculated for Part A, Part B and Part B post-flare separately. Part B summaries will not consider any post-flare escape treatment injection. 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. The sum of the difference in number of syringes between the actual used and expected syringes will be summarized. In addition, a ratio of compliance will be further computed based on the number of actual and expected syringes. The ratio of compliance will be summarized as a continuous variable and categorically (<0.80, ≥0.80–≤1.0 and >1.0). The general formula for the compliance ratio (CR) is given as follows:

CR = # actual syringes / # expected syringes

The second approach defines compliance with study drug administration based upon comparing the actual day of administration with the expected day of administration. The expected day of administration will be based upon (1) the Baseline date and (2) the previous injection date. The

sum of the absolute difference in days between the actual and expected days will be summarized. In addition, a ratio of compliance will be computed based upon the actual and expected days for each of the 2 methods. The ratio of compliance will be summarized as a continuous variable and categorically (<0.80, ≥0.80). The general formula for the compliance ratio is given as follows:

Compliance Ratio (CR) = {[Exposure Duration (days)] - [Cumulative Difference (days)]}/ {Exposure Duration (days)}

The CR ranges between 0 and 1. Specific details are provided in Section 12.2.

Compliance based on these methods will be summarized by axSpA subpopulation and overall in Part A using the SS, by treatment group in Part B using the SSB and by treatment group in Part B Post-flare using the ETS. The compliance for Part B will be calculated based on the first injection in Part B and the compliance for Part B post-flare will be calculated based on the first injection in Escape Part.

If the Week 48 injection was not performed, the date of the Week 48 visit will be used as reference point for calculating the compliance of the first injection in Part B. Injections performed after the first injection in Part B will use the date of first injection in Part B as reference point for calculating the compliance of these injections.

If no Escape Week 0 injection was performed, the date of the Escape Week 0 visit will be used as reference point for calculating the compliance of the first injection in Escape Part. Injections performed after the first injection in Escape Part will use the date of first injection in Escape Part as reference point for calculating the compliance of these injections.

A by-subject listing of compliance will be provided by treatment group and axSpA subpopulation based on the SS.

8 EFFICACY ANALYSES

Efficacy summaries for variables collected in Part A will be based on the OLS. All efficacy analyses for Part B will be performed using the RS, unless otherwise mentioned (ie, the FAS and PPS will be used for a sensitivity analysis on the primary endpoint only). Efficacy summaries for subjects experiencing a flare in Part B will be performed using the FS. All efficacy endpoints will be listed using the observed case data. Furthermore, for selected analyses, statistical appendices will be created and included in the CSR appendix “Documentation of Statistical Methods”.

8.1 Primary efficacy variable

8.1.1 Primary analysis of the primary efficacy variable

The number and percentage of subjects in Part B who do not experience a flare will be summarized descriptively for the RS using NRI data. The 95% CI based on the Exact Binomial distribution will also be presented. Summary statistics will be presented by treatment group in Part B.

The primary analysis will be based on a logistic regression model which will include factors of treatment group, region, and mNY classification. The final values of the randomization stratification factors (region, mNY classification) will be used for this analysis. The odds ratios based on the proportion of subjects who do not experience a flare will be estimated from this

logistic regression model and presented with 95% 2-sided CIs and p-values. This analysis will be conducted for the RS using NRI data.

To avoid the problem of the monotone likelihood resulting in infinite large confidence intervals (eg, if one of the cell counts in the 2x2 table is equal to zero), a penalized maximum likelihood approach based on the modified score procedure of Firth (Heinze and Schemper, 2002) will be used in the logistic models.

Each CZP dose will be compared against PBO to establish superiority over PBO and will be tested sequentially at an alpha of 0.05. The primary statistical analysis of the primary efficacy variable will account for the testing of multiple doses 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 the given CZP dose and PBO, will be performed in the sequence shown below:

1. CZP 200mg Q2W vs PBO
2. CZP 200mg Q4W vs PBO

The second test will be performed irrespective of whether the first test is significant at the 0.05 level or not. However, the second test will be interpreted as statistically significant only if the first test is significant at the 0.05 level as well.

8.1.2 Secondary analyses of the primary efficacy variable

The number and percentage of subjects in Part B who do not experience a flare will be summarized for the following subgroups: age, gender, race, region, baseline CRP level, overall ADAb status and mNY classification. Factors for subgroups and the interaction term of subgroup and treatment will be added to the primary model specified in Section 8.1.1. The 95% CI based on the Exact Binomial distribution will also be presented. These summaries will be performed for the RS using NRI data and will be based on descriptive statistics only.

8.1.3 Supportive and sensitivity analyses of the primary efficacy variable

The number and percentage of subjects in Part B who do experience a flare will be summarized descriptively for the RS using NRI data, indicating the contributing event (ASDAS defined flare using scheduled visits only or using unscheduled visits, missing ASDAS values, premature study discontinuation with reason “AE or Lack of efficacy” or other reason, incorrectly ‘flared’ due to delay in data entry).

The flare by IXRS, calculated flare, time to flare and reason for flare according to NRI analysis will be listed based on the RS.

The analysis described in Section 8.1.1 will be repeated for the RS using the different approaches as defined in Section 3.2.22 (observed case, premature withdrawal, MMRM imputation, referenced-based multiple imputation). The 95% CI based on the Exact Binomial distribution will also be presented.

In addition, the analysis as described in Section 8.1.1 will also be repeated for the FAS and PPS using NRI data.

8.2 Secondary efficacy variables

8.2.1 Part A

All summaries of secondary efficacy variables in Part A will be based on the OLS.

8.2.1.1 Percentage of subjects achieving sustained remission at Week 48

The number and percentage of subjects achieving sustained remission at Week 48 will be summarized in frequency tables by axSpA subpopulation and for all subjects combined. The 95% CI based on the Exact Binomial distribution will also be presented. The presentation will be done on the observed case data and will be repeated on the imputed data using NRI.

8.2.1.2 ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD and ASDAS-vHD) at Week 48

ASDAS disease activity at Week 48 will be summarized in frequency tables by axSpA subpopulation and for all subjects combined. The number and percentage of subjects in each of the 4 disease activity levels will be presented.

The presentation will be done using observed case data and will be repeated using LOCF imputation for the missing data.

8.2.1.3 ASDAS clinical improvement (ASDAS-MI and ASDAS-CII) at Week 48

The number and percentage of subjects achieving ASDAS-MI and ASDAS-CII at Week 48 will be summarized in frequency tables by axSpA subpopulation and for all subjects combined. The presentation will be done using observed case data and will be repeated using NRI data.

8.2.2 Part B

The following analyses are not part of the multiplicity-controlled testing procedure described in Section 8.1. The p-values reported for these analyses will not be adjusted for multiplicity and will be considered nominal. All analyses and summaries of the efficacy variables in Part B will be based on the RS. All tables will be presented by treatment group in Part B.

8.2.2.1 Time to flare

The time to flare will be analyzed using Kaplan-Meier methods, comparing each CZP dose to PBO. Stratified two-sided log rank test using strata geographic region (North America, Western Europe, Eastern Europe and Asia) and mNY classification (mNY positive, mNY negative) will be used to evaluate between-group differences.

Kaplan-Meier plots of the time to flare will also be produced. Kaplan-Meier estimates of percentages of subjects experiencing a flare by Week 52, Week 56, Week 60, Week 64, Week 68, Week 72, Week 76, Week 80, Week 84, Week 88, Week 92 and Week 96 will be provided along with the corresponding 95% CI by treatment arm. For the calculation of the 95% CI, the LOGLOG transformation will be used.

These analyses will be repeated using the different approaches as defined in Section 3.2.23 (observed case, premature withdrawal).

8.2.2.2 Continuous secondary efficacy variables

Summary statistics of observed and change from Part B Baseline to Week 96 values for ASDAS, BASDAI, BASFI, and BASMI will be presented by treatment group. These summary statistics will be repeated for Part A Baseline.

The changes from Part B Baseline to Week 96 in ASDAS, BASDAI, BASFI and BASMI will be compared between treatment groups using a MMRM. The pattern of missingness for these variables is assumed to be MAR. Post-flare data for subjects experiencing a flare (in any treatment group) will be treated as missing in the MMRM analysis. The MMRM methods will be applied to all observed change from Part B Baseline data obtained from Week 48 up to Week 96. The model will include Part B Baseline score as a fixed-effect covariate, treatment group, region, mNY classification, and visit as fixed-effect categorical factors, and Part B Baseline-by-visit and treatment group-by-visit as interaction terms. The multiple visits for each subject will be incorporated as repeated measures (fixed term) within each subject (random term). An unstructured covariance matrix will be used and visit will be treated as a categorical predictor. An unstructured correlation pattern will be used to estimate the variance-covariance of the within-subject repeated measures. Summary tables will present the adjusted means for PBO and the CZP doses, the respective differences to PBO, the corresponding p-values, and 95% CIs.

The comparison between treatment groups using MMRM will be repeated for Part A Baseline with Part A Baseline score as a fixed-effect covariate, treatment group, region, mNY classification, and visit as fixed-effect categorical factors, and Part A Baseline-by-visit and treatment group-by-visit as interaction terms.

Summary statistics of observed and change from Part B Baseline to Week 96 values for SIJ SPARCC and Spine ASspiMRI-a in the Berlin modification scores will be presented by treatment group. These summaries will be based on observed case data.

The changes from Part B Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores at Week 96 will be compared between treatment groups using an analysis of covariance (ANCOVA) model. The model will include Part B Baseline score, treatment group, region, and mNY classification. Summary tables will present the adjusted means for PBO and the CZP doses, the respective differences to PBO, the corresponding p-values, and 95% CIs. This analysis will be performed using observed case data only.

These summary statistics will be repeated for Part A Baseline and the model includes Part A Baseline score, treatment group, region, and mNY classification.

8.2.2.3 Binary secondary efficacy variables

The number and percentage of subjects in the ASDAS disease activity levels (ASDAS-ID, ASDAS-LD, ASDAS-HD, ASDAS-vHD) as well as the number and percentage of subjects achieving ASDAS clinical improvement (ASDAS-CII, ASDAS-MI), ASAS response (ASAS20, ASAS40, ASAS5/6, ASAS PR) and BASDAI50 response will be summarized at Week 96 by treatment group. These summaries will be based on observed case data and will be repeated using NRI data (for all variables except ASDAS disease activity level); improvement will be relative to Part A Baseline only.

The statistical analysis of ASDAS clinical improvement (ASDAS-CII, ASDAS-MI), ASAS response (ASAS20, ASAS40, ASAS5/6, ASAS PR) and BASDAI50 response at Week 96 will

be based on a logistic regression model which will include factors of treatment group, region, and mNY classification. The responder rates for each treatment group and the adjusted odds ratios with reference to PBO will be estimated from this logistic regression model and presented with 95% 2-sided CIs and p-values. This analysis will be performed using NRI data.

8.2.3 Subjects experiencing a flare in Part B

All summaries will be based on observed case data and will use the FS, ie only subjects experiencing a flare based on the IXRS data will be included in these summaries. If deemed needed, further sensitivity analyses may be conducted excluding subjects incorrectly 'flared'.

8.2.3.1 Continuous secondary efficacy variables

For the continuous variables (ASDAS, BASDAI, BASFI, BASMI, SIJ SPARCC score and spine ASspiMRI-a in the Berlin modification score), descriptive statistics by randomized treatment group will be displayed. Summary tables will include the actual results at Week 96 (or later), as well as change from Part B Baseline.

These summary statistics will be repeated for Flare Baseline.

8.2.3.2 Binary secondary efficacy variables

The number and percentage of subjects in the ASDAS disease activity levels (ASDAS-ID, ASDAS-LD, ASDAS-HD, ASDAS-vHD) as well as the number and percentage of subjects achieving ASDAS clinical improvement (ASDAS-CII, ASDAS-MI), and ASAS response (ASAS20, ASAS40, ASAS5/6, and ASAS PR) will be summarized at the Week 96 (or later) visit by randomized treatment group.

8.3 Other efficacy variables

8.3.1 Part A

8.3.1.1 ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD and ASDAS-vHD) and clinical improvement (ASDAS-CII, ASDAS-MI)

The number and percentage of subjects in each of the 4 ASDAS disease activity levels will be presented for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data and will be repeated using LOCF imputation for missing data.

The number and percentage of subjects achieving ASDAS-MI and ASDAS-CII will be summarized in frequency tables for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data and will be repeated using NRI data.

8.3.1.2 ASAS20, ASAS40, ASAS5/6, and ASAS PR response

Frequency tables will be produced to show the number and percentage of responders for each ASAS endpoint for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data and will be repeated using NRI data.

8.3.1.3 ASDAS, BASDAI, and BASMI

Summary statistics of actual values and change from Part A Baseline values will used to summarize each score (ASDAS, BASDAI, and BASMI) for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data and will be repeated using LOCF imputation for missing data.

8.3.1.4 SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the SIJ SPARCC and Spine ASspiMRI-a in the Berlin modification scores for each visit by axSpA subpopulation and for all subjects combined. The analysis will use results of the single read assessments for Part A from the independent reviewer. These summaries will be produced for the OLS using observed case data only.

8.3.1.5 Individual ASAS core components

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the ASAS core components for each visit by axSpA subpopulation and for all subjects combined:

- PtGADA
- Total spinal pain (NRS)
- BASFI
- Average of Questions 5 and 6 of the BASDAI concerning morning stiffness

These summaries will be produced for the OLS using observed case data and will be repeated for BASFI and PtGADA using LOCF imputation for missing data.

8.3.1.6 PhGADA

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize PhGADA for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data.

8.3.1.7 Nocturnal spinal pain (NRS)

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize nocturnal spinal pain for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data and will be repeated using LOCF imputation for missing data.

8.3.1.8 Fatigue (NRS) (from BASDAI)

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize fatigue item score of the BASDAI for each visit by axSpA subpopulation and for all subjects combined. This summary will be produced for the OLS using observed case data and will be repeated using LOCF imputation for missing data.

The remaining components of the BASDAI maybe analyzed as well.

8.3.1.9 BASDAI50 response

Frequency tables will be produced to show the number and percentage of BASDAI50 responders for each visit by axSpA subpopulation and for all subjects combined. This summary will be performed on the OLS using observed case data and will be repeated using NRI data.

8.3.1.10 C-reactive protein (CRP)

Summary statistics of the actual values, change from Part A Baseline values and the ratio of post-Baseline over Part A Baseline values will be used to summarize the CRP values for each visit by axSpA subpopulation and for all subjects combined. This summary will be produced for the OLS using observed case data and will be repeated using LOCF imputation for missing data.

The geometric mean and the corresponding CV of the geometric mean will be computed and displayed for the observed values and for the ratio of post-Baseline over Part A Baseline values.

8.3.1.11 Ankylosing Spondylitis Quality of Life (ASQoL)

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the ASQoL score for each visit by axSpA subpopulation and for all subjects combined. This summary will be produced for the OLS using observed case data.

8.3.1.12 Work Productivity Survey (WPS)

Frequency tables will be produced to summarize the answer provided to question 1 of the WPS for each visit by axSpA subpopulation and for all subjects combined.

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the answers to the question 2 to 9 of the WPS for each visit by axSpA subpopulation and for all subjects combined.

These summaries will be produced for the OLS using observed case.

8.3.1.13 EuroQoL Health Status Questionnaire 5 dimensions 3 levels (EQ-5D-3L)

Frequency tables will be produced to summarize answers provided to each of the 5 items of the EQ-5D-3L for each visit by axSpA subpopulation and for all subjects combined.

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the EQ-5D-3L VAS score for each visit by axSpA subpopulation and for all subjects combined.

These summaries will be produced for the OLS using observed case data.

8.3.1.14 Maastricht Ankylosis Spondylitis Enthesitis Score (MASES)

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the MASES score for each visit by axSpA subpopulation and for all subjects combined. This summary will be produced for the OLS using observed case data.

8.3.1.15 Swollen and tender joint counts (44 joint count)

Summary statistics of the actual and change from Part A Baseline values will be used to summarize TJC and SJC for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data.

8.3.1.16 Spinal mobility

Summary statistics of the actual and change from Part A Baseline values will be used to summarize spinal mobility as assessed by occiput to wall distance and spinal mobility as assessed by chest expansion for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data.

8.3.1.17 Number of uveitis flares, IBD exacerbations and psoriasis exacerbations

Separate tables will be produced to summarize incidence of each of these events (uveitis flares, IBD exacerbations and psoriasis exacerbations) during Part A.

These summary tables will display the number and percentage of subjects experiencing the considered event in Part A and number of event. In this style of output, "number of event" will include all occurrences of the event including repeat occurrences in individual subjects, while "number of subjects" will count each subject only once.

The presentation will be based on the OLS by axSpA subpopulation and overall using observed case data.

For each of these events (uveitis flares, IBD exacerbations and psoriasis exacerbations), the summary table will be repeated based on the subset of subjects with a current occurrence of the event at screening.

8.3.1.18 Fecal calprotectin

Actual fecal calprotectin levels and corresponding change from Part A Baseline values will be summarized by axSpA subpopulation and overall for each visit at which stool samples were taken during Part A using the geometric mean, 95% CI, geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. This analysis will be performed on observed case data only using the OLS.

8.3.1.19 Serum calprotectin

Actual serum calprotectin levels and corresponding change from Part A Baseline values will be summarized by axSpA subpopulation and overall for each visit at which stool samples were taken during Part A using the geometric mean, 95% CI, geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. This analysis will be performed on observed case data only using the OLS.

8.3.1.20 Inflammatory Bowel Disease Questionnaire (IBD-Q)

Summary statistics will be used to summarize each score (Global IBD-Q total score, IBD-Q bowel symptoms score, IBD-Q systemic symptoms score, IBD-Q emotional function score and IBD-Q social function score) using the actual values and change from Part A Baseline values at each visit by axSpA subpopulation and overall. These summaries will be produced for the OLS using observed case data.

Frequency tables by axSpA subpopulation and overall will be produced to show the number and percentage of IBD-Q responders and IBD-Q remitters at each visit. These summaries will be produced for the OLS using observed case data and will be repeated using NRI data.

8.3.1.21 Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP

The following correlations will be summarized at Part A Baseline and Week 48 by axSpA subpopulation and overall:

- fecal calprotectin and serum calprotectin
- fecal calprotectin and global IBD-Q total score
- serum calprotectin and global IBD-Q total score
- fecal calprotectin and CRP
- serum calprotectin and CRP
- CRP and global IBD-Q total score

For each considered correlations, number and percentage of subjects with non-missing values and Spearman's rank correlation coefficient along with corresponding 2-sided 95% CI will be reported at each visit. These summaries will be produced for the OLS using observed case data.

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

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize each SF-36 domain scores and the PCS and MCS scores for each visit by axSpA subpopulation and for all subjects combined. These summaries will be produced for the OLS using observed case data.

Frequency tables will be produced to show the number and percentage of SF-36 PCS and MCS responders for each visit by axSpA subpopulation and for all subjects combined. This summary will be performed on the OLS using observed case data and will be repeated using NRI data.

8.3.1.23 Resources utilization

Summary statistics and frequency distribution (number of subjects by number of medical resources used) will be presented by axSpA subpopulation and all subjects for concomitant medical procedures, health care provider consultations, and in-patient hospitalizations and emergency room visits, including length of stay. These summaries will be produced for the OLS using observed case data.

8.3.2 Part B

The following analyses are not part of the multiplicity-controlled testing procedure described in Section 8.1.1. The p-values reported for these analyses will not be adjusted for multiplicity and will be considered nominal. All efficacy summaries and analyses for Part B will be performed using the RS.

8.3.2.1 ASDAS clinical improvement (ASDAS-CII, ASDAS-MI)

The number and percentage of subjects achieving ASDAS-MI and ASDAS-CII will be summarized in frequency tables by visit and by treatment group. This summary will be produced using observed case data and will be repeated using NRI data. This analysis will be performed using Part A Baseline as reference for the definition of ASDAS-MI and ASDAS-CII.

The statistical analysis of ASDAS-MI and ASDAS-CII at each visit will be based on a logistic regression model which will include factors of treatment group, region, and mNY classification.

The responder rates and adjusted odds ratios with reference to PBO will be estimated from this logistic regression model and presented with 95% 2-sided CIs and p-values. These analyses will be performed using NRI data. These analyses will be performed using Part A Baseline as reference for the definition of ASDAS-MI and ASDAS-CII.

8.3.2.2 ASAS20, ASAS40, ASAS5/6, and ASAS PR response

The number and percentage of subjects achieving each of the ASAS endpoints will be summarized in frequency tables by visit and by treatment group. This summary will be produced using observed case data and will be repeated using NRI data. This analysis will be performed using Part A Baseline as reference for the definition of the ASAS endpoints.

The statistical analysis of ASAS20, ASAS40, ASAS5/6 and ASAS PR at each visit will be based on a logistic regression model which will include factors of treatment group, region, and mNY classification. The responder rates and adjusted odds ratios with reference to PBO will be estimated from this logistic regression model and presented with 95% 2-sided CIs and p-values. These analyses will be performed using NRI data. These analyses will be performed using Part A Baseline as reference for the definition of the ASAS endpoints.

8.3.2.3 ASDAS, BASDAI, BASFI and BASMI

Summary statistics of the actual values, change from Part B Baseline values and change from Part B Baseline values for ASDAS, BASDAI, BASFI and BASMI will be presented by visit and treatment group. These summaries will be produced using observed case data. These summary statistics will be repeated for Part A Baseline.

The changes from Part B Baseline to each visit in Part B for ASDAS, BASDAI, BASFI and BASMI will be compared between treatment groups using a MMRM. The model will include Part B Baseline score as a fixed-effect covariate, treatment group, region, mNY classification, and visit as fixed-effect categorical factors, and Part B Baseline-by-visit and treatment group-by-visit as interaction terms. The multiple visits for each subject will be incorporated as repeated measures (fixed term) within each subject (random term). An unstructured covariance matrix will be used and visit will be treated as a categorical predictor. Summary tables will present the adjusted means for PBO and the CZP doses, the respective differences to PBO, the corresponding p-values, and 95% CIs.

This analysis will be repeated based on change from Part A Baseline with Part A Baseline score as a fixed-effect covariate, treatment group, region, mNY classification, and visit as fixed-effect categorical factors, and Part A Baseline-by-visit and treatment group-by-visit as interaction terms.

8.3.2.4 Individual ASAS core components

Summary statistics will be used to summarize the actual values, change from Part A Baseline values and change from Part B Baseline values by visit for the ASAS score components (note that BASFI is covered in Section 8.3.2.3):

- PtGADA
- Total spinal pain (NRS)
- Average of Questions 5 and 6 of the BASDAI concerning morning stiffness

These summaries will be produced using observed case data.

For each component, the changes from Part B Baseline to each visit in Part B will be compared between treatment groups using a MMRM. (see Section 8.3.2.3). This analysis will be repeated based on change from Part A Baseline.

8.3.2.5 Fatigue NRS

Fatigue (NRS) will be taken from BASDAI Question 1. Summary statistics for actual values, change from Part A Baseline values and change from Part B Baseline values by visit and by treatment group will be used to summarize the fatigue scores at each visit during Part B. This summary will be produced using observed case.

Treatment group comparisons for both CZP groups vs. PBO will be performed based on the change from Part B Baseline to each visit in Fatigue using a MMRM. (see Section 8.3.2.3). This analysis will be repeated based on change from Part A Baseline.

The remaining components of the BASDAI maybe analyzed as well.

8.3.2.6 BASDAI50 response

Frequency tables by visit and by treatment group will be produced to show the number and percentage of BASDAI50 responders at each visit during Part B. This summary will be produced using observed case data and will be repeated using NRI for missing data. This analysis will be performed using Part A Baseline as reference for the definition of BASDAI50.

The statistical analysis of BASDAI50 response at each visit will be based on a logistic regression model which will include factors of treatment group, region, and mNY classification. The responder rates for each treatment group and the adjusted odds ratios with reference to PBO will be estimated from this logistic regression model and presented with 95% 2-sided CIs and p-values. This analysis will be performed using NRI data. This analysis will be performed using Part A Baseline as reference for the definition of BASDAI50.

8.3.2.7 ASQoL

Summary statistics for actual values, change from Part A Baseline values and change from Part B Baseline values by visit and by treatment group will be used to summarize the ASQoL scores at each visit during Part B. This summary will be produced using observed case.

Treatment group comparisons for both CZP groups vs. PBO will be performed based on the change from Part B Baseline to each visit in ASQoL score using MMRM (see Section 8.3.2.3).

This analysis will be repeated based on change from Part A Baseline.

8.3.2.8 Work Productivity Survey (WPS)

Frequency tables will be produced by visit and by treatment group for the answer provided to question 1 of the WPS.

Summary statistics will be used to summarize the actual values, change from Part A Baseline values and change from Part B Baseline values by visit and by treatment group for answers to questions 2 to 9 of WPS. These summaries will be produced using observed case data.

Treatment group comparisons for both CZP groups vs. PBO will be performed for the change from Part B Baseline for answers to questions 2 to 9 using MMRM (see Section 8.3.2.3).

This analysis will be repeated based on the change from Part A Baseline.

8.3.2.9 Number of uveitis flares, IBD exacerbations and psoriasis exacerbations

Separate tables will be produced to summarize incidence of each of these events (uveitis flares, IBD exacerbations and psoriasis exacerbations) during Part B.

These summary tables will display the number and percentage of subjects experiencing the considered event in Part B, number of event, incidence rate, event rate and 100 patient exposure year. Only the events occurring up to Escape Week 0 will be considered. In this style of output, "number of event" will include all occurrences of the event including repeat occurrences in individual subjects, while "number of subjects" will count each subject only once.

The presentation will be performed by treatment group using observed case data.

For each of these events (uveitis flares, IBD exacerbations and psoriasis exacerbations), the summary table will be repeated based on the subset of subjects with a history of occurrence and/or a current occurrence of the event at screening.

8.3.2.10 Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP

The following correlations will be summarized at Week 48 and Week 96 by treatment group:

- fecal calprotectin and serum calprotectin
- fecal calprotectin and global IBD-Q total score
- serum calprotectin and global IBD-Q total score
- fecal calprotectin and CRP
- serum calprotectin and CRP
- CRP and global IBD-Q total score

For each considered correlations, number and percentage of subjects with non-missing values and Spearman's rank correlation coefficient along with corresponding 2-sided 95% CI will be reported at each visit. These summaries will be produced using observed case data only.

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

Summary statistics for actual values, change from Part A Baseline values and change from Part B Baseline values by visit and by treatment group will be used to summarize the SF-36 domain, PCS and MCS scores at each visit during Part B. This summary will be produced using observed case data.

Treatment group comparisons for both CZP groups vs. PBO will be performed based on the change from Part B Baseline in the SF-36 physical function domain, PCS and MCS scores only at the time points of Part B using MMRM (Section 8.3.2.3). This analysis will be repeated based on change from Part A Baseline.

Frequency tables by visit and by treatment group will be produced to show the number and percentage of SF-36 PCS and MCS responders at each visit during Part B. This summary will be produced using observed case data and will be repeated using NRI for missing data. This analysis will be performed using Part A Baseline as reference for the definition of SF-36 PCS and MCS responders.

The statistical analysis of SF-36 PCS and MCS response at each visit will be based on a logistic regression model which will include factors of treatment group, region, and mNY classification. The responder rates for each treatment group and the adjusted odds ratios with reference to PBO will be estimated from this logistic regression model and presented with 95% 2-sided CIs and p-values. This analysis will be performed using NRI data. This analysis will be performed using Part A Baseline as reference for the definition of SF-36 PCS and MCS response.

8.3.2.12 Descriptive summaries of other efficacy variables

The below listed efficacy variables will be summarized descriptively by visit and by treatment group as applicable using observed case based on the RS unless otherwise specified.

- ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD and ASDAS-vHD)
- Change from Part A Baseline and change from Part B Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores (summaries based on observed case data only)
- Change from Part A Baseline and change from Part B Baseline in PhGADA
- Change from Part A Baseline and change from Part B Baseline in nocturnal spinal pain (NRS)
- CRP (actual values, change from Part A Baseline values, change from Part B Baseline values, ratio of post-Baseline over Part A Baseline values, ratio of post-Baseline over Part B Baseline values)
- Health status as assessed by the EQ-5D-3L: domains, VAS actual score, change from Part A Baseline and change from Part B Baseline in VAS score
- Change from Part A Baseline and change from Part B Baseline in MASES
- Change from Part A Baseline and change from Part B Baseline in swollen and tender joint counts (44 joint counts). The summaries for TJC will be repeated including only the subjects with a TJC>0 at Part A Baseline or Part B Baseline, as appropriate. Similar approach will be used for SJC.
- Change from Part A Baseline and change from Part B Baseline in spinal mobility as assessed by occiput to wall distance and in spinal mobility as assessed by chest expansion
- Change from Part A Baseline and change from Part B Baseline in fecal calprotectin (summaries based on observed case data only)
- Change from Part A Baseline and change from Part B Baseline in serum calprotectin (summaries based on observed case data only)
- Change from Part A Baseline and change from Part B Baseline in the IBD-Q scores (Global IBD-Q total score, IBD-Q bowel symptoms score, IBD-Q systemic symptoms score, IBD-Q emotional function score and IBD-Q social function score)
- Frequency of IBD-Q responders and IBD-Q remitters (NRI data will be used) using Part A Baseline and Part B Baseline as reference
- Resources utilization: concomitant medical procedures, healthcare provider consultations not foreseen by the protocol, in-patient hospitalizations and emergency room visits

8.3.3 Subjects experiencing a flare in Part B

The below listed efficacy variables for subjects experiencing a flare in Part B will be summarized descriptively by visit and randomized treatment group using the FS. If deemed needed, further sensitivity analyses may be conducted excluding subjects incorrectly 'flared'. All summaries will be based on observed case data and will present data as collected at all scheduled study visits following Escape Week 0. Generally, the format of the summaries described here will mirror the format of summaries described in Section 8.3.2.

- ASDAS disease activity (ASDAS-ID, ASDAS-LD, ASDAS-HD, and ASDAS-vHD)
- ASDAS clinical improvement (ASDAS-CII, ASDAS-MI, MI) using Flare Baseline as reference
- ASAS20, ASAS40, ASAS5/6, and ASAS PR using Flare Baseline as reference
- Change from Part B Baseline and change from Flare Baseline in ASDAS, BASDAI and BASMI
- Change Part B Baseline and change from Flare Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores
- Change from Part B Baseline and change from Flare Baseline in all individual ASAS core components:
  - PtGADA
  - Total spinal pain (NRS)
  - BASFI
  - Average of questions 5 and 6 of the BASDAI concerning morning stiffness
- Change from Part B Baseline and change from Flare Baseline in PhGADA
- Change from Part B Baseline and change from Flare Baseline in nocturnal spinal pain (NRS)
- Change from Part B Baseline and change from Flare Baseline in Fatigue (NRS) (from BASDAI). The remaining components of the BASDAI maybe analyzed as well.
- BASDAI50 response using Flare Baseline as reference
- CRP (actual values, change from Part B Baseline values, ratio of post-Baseline over Part B Baseline values, change from Flare Baseline values, ratio of post-Baseline over Flare Baseline values)
- Change from Part B Baseline and change from Flare Baseline in ASQoL
- Change from Part B Baseline and change from Flare Baseline WPS
- Health status as assessed by the EQ-5D-3L: domains, VAS actual score, change from Part B Baseline and change from Flare Baseline in VAS score
- Change from Part B Baseline and change from Flare Baseline in MASES
- Change from Part B Baseline and change from Flare Baseline in spinal mobility as assessed by occiput to wall distance and in spinal mobility as assessed by chest expansion

- Number of uveitis flares, IBD exacerbations and psoriasis exacerbations occurring post ASDAS-defined flare
- Change from Part B Baseline and change from Flare Baseline in fecal calprotectin
- Change from Part B Baseline and change from Flare Baseline in serum calprotectin
- Change from Part B Baseline and change from Flare Baseline in the IBD-Q scores (Global IBD-Q total score, IBD-Q bowel symptoms score, IBD-Q systemic symptoms score, IBD-Q emotional function score and, IBD-Q social function score)
- Frequency of IBD-Q responders and IBD-Q remitters using Part B Baseline and Flare Baseline as reference
- Correlation between fecal calprotectin, serum calprotectin, IBD-Q and CRP
- Change from Part B Baseline and change from Flare Baseline in all SF-36 domains, SF-36 PCS, SF-36 MCS
- Frequency of SF-36 PCS responders and SF-36 MCS responders using Flare Baseline as reference
- Resources utilization: concomitant medical procedures, healthcare provider consultations not foreseen by the protocol, in-patient hospitalizations and emergency room visits

9 PHARMACOKINETICS AND PHARMACODYNAMICS

9.1 Pharmacokinetics

CZP plasma concentrations will be tabulated and summarized using PKSB for each visit at which samples were taken during Part A by axSpA subpopulation using the geometric mean, 95% CI, geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. Plasma concentration time curves of Part A will be plotted for all subjects and Part A treatment-emergent ADAb (Part A TE ADAb) status with titer classification (see Section 9.2).

CZP plasma concentrations will be tabulated and summarized by treatment group using PKSB for each visit of the total studyat which samples were taken using the geometric mean, 95% CI, geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. The table will be repeated for TE ADAb status with titer classification (see Section 9.2.).

For each treatment group, plasma concentration time curves of Part B will be plotted and repeated for TE ADAb status with titer classification (see Section 9.2.).

Part B figures will not include post-flare assessments. Separate figures will be produced for subjects experiencing a flare in Part B and start escape therapy including only the post-flare assessments.

Spaghetti plots of CZP plasma concentrations by week from CZP first dosing separated by treatment group and TE ADAb status and titer classification (see Section 9.2.) will be presented for subjects with flare and no flare per NRI approach.

Spaghetti plots of CZP plasma concentrations by week from CZP first dosing separated by treatment group will be presented for flared subjects per NRI by TE ADAb status and titer classification (see Section 9.2.).

Individual plots by subject of CZP Concentrations/ADAb titer and ASDAS score plotted on the Y-axes by visit (X-axis) for the full treatment period, including SFU. Flare data will be displayed on the days relative to Part A.

Concentrations below the limit of quantification (BLQ) will be replaced by the LLOQ divided by 2 for the calculation of descriptive statistics.

Anomalous values will be not included in summaries/analysis and will be reviewed and flagged by pharmacokineticist.

All CZP plasma concentration will be listed at subject-level.

9.2 Immunogenicity

Immunogenicity will be assessed through listing of individual results by subject and summary tables. Immunogenicity data will be correlated with PK and efficacy readout.

A cut point will be determined by the bioanalytical laboratory during assay validation. This cut point will be used to determine the status of ADAb in the test sample.

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 positive screen and positive immunodepletion.
- An ADAb status of negative will be concluded for any sample with an ADAb level that is either negative screen or positive screen and negative immunodepletion.

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 ADAb assay. ADAb titer dilution factor is presented in the listings and summaries excluding the minimum required dilution, i.e. reportable values are divided by 100.

If the titer for an ADAb level that is positive screen and positive immunodepletion is missing, then a conservative will be used and ADAb status will be consider as positive. No imputation rules apply for the missing titer.

If the ADAb level is positive screen but no confirmatory result could be determined, then then a conservative will be used and ADAb status will be consider as positive.

Anomalous value will be not included in summaries/analysis and will be reviewed and flagged by pharmacokineticist.

Subjects will have the following Baseline ADAb status:

Baseline ADAb Status positive is defined as having an ADAb confirmed positive inhibition result at baseline, Baseline ADAb Status negative is defined as having negative inhibition result at baseline.

Subjects will receive a treatment-emergent ADAb status, inclusive of Baseline and Post-Baseline results (including SFU), and be classified as follows based on the ADAb assay results:

- TE ADAb status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥minimum significant ratio (MSR) -fold

increase from baseline on CZP treatment. The MSR will be defined during the process of sample analysis and is disease-specific.

- TE ADAb status negative is defined as having no samples either ADAb positive or with values ≥MSR -fold increase from baseline.

Once determined positive, the highest titer during Part A and Part B (including Escape and SFU) is used to categorize the subject.

- Positive <512,
- Positive 512-<=1024,
- Positive >1024.

In addition, for some outputs also the subjects with TE ADAb status negative will be presented.

The TE ADAb will be also defined for the different treatment periods. For each Period the SFU visit will be considered:

- Part A TE ADAb status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline during Part A or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥MSR increase from baseline on CZP treatment during Part A. Once determined positive, the highest titer during Part A is used to categorize the subject.
- Part A TE ADAb status negative is defined as having no samples either ADAb positive or with values ≥MSR increase from baseline during Part A.
- Part B TE ADAb Status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline during Part B or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥MSR increase from baseline on CZP treatment during Part B. Once determined positive, only the highest titer during Part is used to categorize the subject. For subjects with flare in Part B, any assessments after Escape treatment are not included.
- Part B TE ADAb Status negative is defined as having no samples either ADAb positive or with values ≥MSR increase from baseline during Part B (excluding any values up to 8 weeks post first Part B dose).
- Escape TE ADAb Status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline during Escape Part or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥MSR increase from baseline on CZP treatment during Escape Part B. Once determined positive, only the highest titer during Escape is used to categorize the subject.
- Escape TE ADAb Status negative is defined as having no samples either ADAb positive or with values ≥MSR increase from baseline during Escape Part. Once determined positive, only the highest titer during Escape is used to categorize the subject.

Summaries will be done using PKSB.

Summary of shift from Baseline ADAb status with titer classification to TE ADAb Status with titer classification by treatment group and for all subjects will be presented during the entire study.

The shift table will be repeated for

- Part A TE ADAb status with titer classification to Part B TE ADAb status with titer classification by treatment group and for all subjects during Part B.
- Part B TE ADAb status with Part B titer classification to Escape Part TE ADAb status with Escape titer classification by treatment group and for all subjects during Escape Part.

The time to achieving TE ADAb for all subjects will be analyzed based on Kaplan-Meier methods. Subjects will be considered to have an event at the time where treatment-emergent ADAb positive is first achieved during treatment period excluding Baseline/pre-treatment. Subjects classified as treatment-emergent ADAb negative will be censored at the time of last available ADAb result. The median and 95% CI based on the Kaplan-Meier estimation will also be presented. A plot of time to first ADAb positivity will be presented.

Number and percentage of subjects with ADAb titer above the specified cut point ("categorized") at any visit during the treatment period, separated by treatment group will be presented.

A summary of flare responder during Part B (NRI) and ASDAS MI responder at Week 96 (NRI) by treatment group and by highest ADAb titer ("categorized") per subject will be presented during Part A and Part B (excluding Escape). This table will be repeated for Part B (excluding Escape).

A scatter plot of CZP Plasma Concentration and ADAb titer during Part A for all subjects will be presented. The scatter plot will be repeated for Part B by treatment group.

Spaghetti plots of ADAb titer on a loglog scale by week from CZP first dosing separated by treatment group and TE ADAb status/titer classification will be presented for subjects with flare and no flare per NRI approach.

All individual ADAb results will be listed at subject level.

10 SAFETY ANALYSES

10.1 Extent of exposure

10.1.1 Part A

Exposure to study medication in Part A will be evaluated by summarizing the total number of CZP injections, the total amount of CZP received in mg, the total drug exposure duration in days, the patient time at risk and the total patient years at risk. Summaries will be based on the SS and presented by axSpA subpopulation and overall.

Duration of exposure to study medication in Part A will be calculated (in days) as:

- The date of the last Part A injection – date of first Part A injection + 14 days.

If the first study medication administration in Part B, final contact date or death occurs prior to the completion of the 14-day period, exposure to study medication in Part A will be censored at the earliest of these events.

Patient time at risk in Part A will be calculated (in days) as:

- The date of the last injection in Part A – date of first injection in Part A + 70 days

If the first study medication administration in Part B, final contact date or death occurs prior to the completion of the 70-day period, patient time at risk in Part A will be censored at the earliest of these events.

Total patient years at risk will be calculated as the sum of all the relevant patient time at risk (in days) divided by 365.25.

10.1.2 Part B

10.1.2.1 All subjects in SSB – Prior to Flare

Exposure to study medication in Part B will be evaluated by summarizing the total number of CZP or PBO injections, the total number of CZP injections, the total amount of CZP received in mg, the total drug exposure duration in days, the patient time at risk and the total patient years at risk. Summaries will be based on the SSB and presented by treatment group.

Duration of exposure to study medication in Part B will be calculated (in days) for subjects who do not experience a flare as:

- The date of the last CZP 200mg or PBO Part B injection – date of Week 48 injection + 14 days for subjects in the 200mg CZP Q2W and PBO groups [1]
- The date of the last CZP 200mg Part B injection – date of Week 48 injection + 28 days for subjects in the 200mg CZP Q4W group [2]

Duration of exposure to study medication in Part B will be calculated (in days) for subjects who do experience a flare as:

- The date of the last CZP 200mg or PBO Part B injection prior to start of escape full-dose CZP – date of Week 48 injection + 14 days for subjects in the 200mg CZP Q2W and PBO groups [3]
- The date of the last CZP 200mg Part B injection prior to start of escape full-dose CZP – date of Week 48 injection + 28 days for subjects in the 200mg CZP Q4W group [4]

If the final contact date or death occurs prior to the completion of the 14-day period for [1] or prior to the completion of the 28-day period for [2], exposure to study medication in Part B will be censored at the earliest of these events. If the start of escape full-dose CZP occurs prior to the completion of the 14-day period for [3] or prior to the completion of the 28-day period for [4], exposure to study medication in Part B will be censored at the date of this event.

Patient time at risk in Part B will be calculated (in days) for subjects who do not experience a flare as:

- The date of the last CZP 200mg or PBO Part B injection – date of Week 48 injection + 70 days for subjects in the 200mg CZP Q2W and PBO groups [1]
- The date of the last CZP 200mg Part B injection – date of Week 48 injection + 70 days for subjects in the 200mg CZP Q4W group [2]

Patient time at risk in Part B will be calculated (in days) for subjects who do experience a flare as:

- The date of the last CZP 200mg or PBO Part B injection prior to start of escape full-dose CZP – date of Week 48 injection + 70 days for subjects in the 200mg CZP Q2W and PBO groups [3]
- The date of the last CZP 200mg Part B injection prior to start of escape full-dose CZP – date of Week 48 injection + 70 days for subjects in the 200mg CZP Q4W group [3]

If the final contact date or death occurs prior to the completion of the 70-day period for [1] or [2], patient time at risk in Part B will be censored at the earliest of these events. If the start of escape full-dose CZP occurs prior to the completion of the 70-day period for [3] and [4], patient time at risk in Part B will be censored at the date of this event.

Total patient years at risk will be calculated as the sum of all the relevant patient time at risk (in days) divided by 365.25.

10.1.2.2 After flare

Exposure to study medication in Part B after flare and escape to full-dose CZP will be evaluated by summarizing the total number of CZP or PBO injections, the total number of CZP injections, the total amount of CZP received in mg, the total drug exposure duration in days, the patient time at risk and the total patient years at risk. Summaries will be based on the ETS and presented by treatment group.

Duration of exposure to study medication in Part B after escape to full-dose CZP will be calculated (in days) as:

- The date of the last escape full-dose CZP injection – date of first escape full-dose CZP injection + 14 days for all subjects

If final contact or death occurs prior to the completion of the 14–day period, the exposure duration will be censored at the date of the earliest event.

Patient time at risk in Part B after escape to full-dose CZP will be calculated (in days) as:

- The date of the last escape full-dose CZP injection – date of first escape full-dose CZP injection + 70 days for all subjects

If final contact or death occurs prior to the completion of the 70–day period, the patient time at risk will be censored at the date of the earliest event.

Total patient years at risk will be calculated as the sum of all the relevant patient time at risk (in days) divided by 365.25.

10.1.2.3 Full-dose CZP in Part B

Exposure to full-dose CZP in Part B will be evaluated by summarizing the total number of CZP injections, the total amount of CZP received in mg, the total drug exposure duration in days, the patient time at risk and the total patient years at risk. Summaries will be based on the FDS. These summaries will be produced presenting a single column “CZP full-dose in Part B (CZP 200mg Q2W)”.

Duration of exposure to full-dose CZP in Part B as double-blind or as escape treatment will be calculated (in days) as:

- The date of the last escape full-dose CZP injection – date of the first escape full-dose CZP injection + 14 days for subjects in the 200mg CZP Q4W and PBO groups who escape to full-dose CZP therapy
- The date of the last escape full-dose CZP injection – date of Week 48 injection + 14 days for subjects in the 200mg CZP Q2W group experiencing a flared
- The date of the last Part B injection – date of Week 48 injection + 14 days for subjects in the 200mg CZP Q2W group who did not flare

If final contact or death occurs prior to the completion of the 14–day period, the exposure duration will be censored at the date of the earliest event.

Patient time at risk in Part B under full-dose CZP therapy will be calculated (in days) as:

- The date of the last escape full-dose CZP injection – date of the first escape full-dose CZP injection + 70 days for subjects in the 200mg CZP Q4W and PBO groups who escape to full-dose CZP therapy
- The date of the last escape full-dose CZP injection – date of Week 48 injection + 70 days for subjects in the 200mg CZP Q2W group experiencing a flared
- The date of the last Part B injection – date of Week 48 injection + 70 days for subjects in the 200mg CZP Q2W group who did not flare

If final contact or death occurs prior to the completion of the 70–day period, the patient time at risk will be censored at the date of the earliest event.

Total patient years at risk will be calculated as the sum of all the relevant patient time at risk (in days) divided by 365.25.

10.1.3 Part A and Part B full-dose CZP therapy

Exposure to full-dose CZP therapy in Part A and/or Part B will be evaluated by summarizing the total number of CZP injections, the total amount of CZP received in mg, the total drug exposure duration in days, the patient time at risk and the total patient years at risk. Summaries will be based on the SS. These summaries will be produced presenting a single column “CZP full-dose in Part A and/or Part B (CZP 200mg Q2W)”.

Duration of exposure to full-dose CZP therapy in Part A and/or Part B will be calculated (in days) as:

- The date of the last Part A full-dose CZP injection – date of first Part A full-dose CZP injection + 14 days for subjects not randomized into Part B. Duration of exposure will be censored at the date of earliest of the following events if any of these events occur prior to the completion of the 14-day period: final contact or death.
- The date of the last Part A full-dose CZP injection – date of first Part A full-dose CZP injection + 14 days for subjects randomized into Part B to the PBO or CZP 200mg Q4W arms who do not experience a flare and do not escape to CZP full-dose therapy. Duration of exposure will be censored at the date of earliest of the following events if any of these events

occur prior to the completion of the 14-day period (first study medication administration in Part B, final contact or death).

- The sum of (date of the last Part A full-dose CZP injection – date of first Part A full-dose CZP injection + 14 days) [1] and (date of the last escape full-dose CZP injection – date of first escape full-dose CZP injection + 14 days) [2] for subjects randomized into Part B to the PBO or CZP 200mg Q4W arms who do experience a flare and do escape to CZP full-dose therapy. If first study medication administration in Part B occurs prior to the completion of the 14 day period for [1], [1] will be censored at the date of this event. If final contact or death occurs prior to the completion of the 14-day period for [2], [2] will be censored at the date of the earliest event.
- The date of the last Part B injection – date of first Part A full-dose CZP injection + 14 days for subjects randomized into Part B to CZP 200mg Q2W arm. If final contact or death occurs prior to the completion of the 14-day period, the exposure will be censored at the date of the earliest event.

Patient time at risk in Part A and or Part B under full-dose CZP therapy will be calculated (in days) as:

- The date of the last Part A full-dose CZP injection – date of first Part A full-dose CZP injection + 70 days for subjects not randomized into Part B. Patient time at risk will be censored at the date of earliest of the following events if any of these events occur prior to the completion of the 70-day period: final contact or death.
- The date of the last Part A full-dose CZP injection – date of first Part A full-dose CZP injection + 70 days for subjects randomized into Part B to the PBO or CZP 200mg Q4W arms who do not experience a flare and do not escape to CZP full-dose therapy. Patient time at risk will be censored at the date of earliest of the following events if any of these events occur prior to the completion of the 70-day period (first study medication administration in Part B, final contact or death).
- The sum of (date of the last Part A full-dose CZP injection – date of first Part A full-dose CZP injection + 70 days) [1] and (date of the last escape full-dose CZP injection – date of first escape full-dose CZP injection + 70 days) [2] for subjects randomized into Part B to the PBO or CZP 200mg Q4W arms who do experience a flare and do escape to CZP full-dose therapy. If first study medication administration in Part B occurs prior to the completion of the 70-day period for [1], [1] will be censored at the date of this event. If final contact or death occurs prior to the completion of the 70-day period for [2], [2] will be censored at the date of the earliest event.
- The date of the last Part B injection – date of first Part A full-dose CZP injection + 70 days for subjects randomized into Part B to CZP 200mg Q2W arm. If final contact or death occurs prior to the completion of the 70-day period, the patient time at risk will be censored at the date of the earliest event.

Total patient years at risk will be calculated as the sum of all the relevant patient time at risk (in days) divided by 365.25.

Exposure data will be listed.

10.2 Adverse events

Pre-Treatment and Post-Treatment AEs

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

Adverse events that start more than 70 days after the last administration of study medication will be defined as post-treatment AEs. No tabular summaries of these events will be produced although they will be included in listings.

Treatment-Emergent AEs (TEAEs)

Treatment emergent AEs (TEAEs) will be defined and summarized in four different ways:

[1] TEAEs in Part A: will be defined as all AEs starting on or after the date of first study medication in Part A and up to 70 days after the last dose of study medication in Part A and before the date of the first study medication in Part B if any.

[2] TEAEs in Part B prior to flare: will be defined as all AEs starting on or after the date of first study medication in Part B and up to 70 days after the last dose of study medication in Part B. AEs occurring more than 70 days after the last dose in Part B will not be considered treatment emergent, regardless of whether this occurs during the SFU period, or after a treatment gap of more than 70 days. AEs which emerge post-flare will be excluded from any summary tables of TEAEs in Part B prior to flare.

[3] TEAEs in Part B in subjects experiencing a flare: Pre-flare TEAEs in Part B will be defined as any TEAEs in Part B with an onset date before the start date of escape therapy. Post-flare TEAEs in Part B will be defined as any TEAEs in Part B with an onset date on or after the start date of escape therapy.

[4] TEAEs in Part B under CZP full-dose therapy: These will be defined for subjects in the FDS who ever received a dose of CZP 200mg Q2W during Part B as double-blind or as escape treatment.

TEAEs in Part B under CZP full-dose therapy will be defined as follows:

- In subjects randomized into Part B to the PBO or CZP 200mg Q4W arm who do experience a flare and do escape to CZP full-dose therapy, TEAEs under CZP full-dose therapy will include all post-flare TEAEs in Part B
- In subjects randomized into Part B to CZP 200mg Q2W arm, TEAEs under CZP full-dose therapy will include all TEAEs in Part B

Data Handling Rules for AEs

Duration of AEs will not be calculated if there is missing stop date information.

If the intensity of an AE is unknown, it will be considered as severe. If the relationship to study drug is missing, it will be considered as related.

For relatedness, AEs determined to be ‘not related’ or ‘unlikely related’ to study drug will map to ‘not related’, while all other values will map to ‘related’.

AE Summaries

All AE summary tables described in this Section will be produced four ways, ie for each of the following types of AEs:

a. TEAEs in Part A, based on the SS

b. TEAEs in Part B prior to flare, based on the SSB

c. TEAEs in Part B for subjects experiencing a flare, based on the ETS

d. TEAEs in Part B under CZP full-dose therapy, based on the FD

SAEs will be classified by SOC, High Level Term (HLT), and PT according to MedDRA®. The latest version available to the sponsor at the time of coding will be used. Summary tables will display the number and percentage of subjects experiencing the TEAEs and number of TEAEs. In this style of output, "number of TEAEs" will include all occurrences of a TEAE including repeat occurrences in individual subjects, while "number of subjects" will count each subject only once.

Overall summaries of AEs will be produced to present the number of events and number and percentage of subjects with:

- Any TEAE
- Any serious TEAE
- Any TEAE leading to study discontinuation
- Permanent withdrawal of study medication due to TEAEs
- Any drug-related TEAE
- Any severe TEAE
- All Deaths (AEs leading to death)
- Deaths (TEAE leading to death)

In addition, the following AE summaries will be presented:

- All TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All serious TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All TEAEs leading to permanent discontinuation of study drug by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All TEAEs leading to death by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c)
- All TEAEs by primary SOC, PT, and maximum intensity (for the Parts/Analysis sets a, b, c, d)
- All TEAEs by primary SOC, PT, and maximum relationship (for the Parts/Analysis sets a, b, c, d)

- All serious TEAEs by primary SOC, PT and maximum relationship (for the Parts/Analysis sets a, b, c, d)
- All TEAEs above reporting threshold of 5% by primary SOC and PT (for the Parts/Analysis sets a, b, c, d)
- All non-serious TEAEs above reporting threshold of 5% by primary SOC and PT (for the Parts/Analysis sets a, b, c)
- All injection related TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)

AE summary tables will also be produced to display the exposure adjusted incidence rate (EAIR) with associated 95% CI, and the exposure adjusted event rate (EAER):

- All TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All serious TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)

For EAIR, the numerator will be the total number of subjects experiencing the AE of interest in the corresponding period. The denominator will be 100 patient-years; that is, the total summation of individual patient-years at risk up to the first occurrence of the AE of interest for subjects with that AE, and the total patient-years at risk for those subjects not experiencing that AE, divided by 100. EAIRs will be presented with a 95% exact CI based upon the Chi-Square distribution (Ulm, 1990).

For EAERs, the numerator will be the number of AEs including repeat occurrences in individual subjects; the denominator will be 100 patient-years. That is, the total summation of individual patient-years at risk divided by 100. No CI will be computed for EAER.

The TEAE summary table will be presented by TE ADAb status and depending on the timepoint of onset of TEAE (Prior to becoming ADAb Positive>=1024, After becoming ADAb Positive>=1024, Subjects Who Never Became ADAb Positive>=1024) based on PKSB.)

The following AEs of interest will be summarized:

1. Serious infections, including opportunistic infections
2. Malignancies, including lymphoma
3. Serious cardiovascular events (i.e. Major adverse cardiac events or MACE)
4. Congestive heart failure
5. Demyelinating-like disorders
6. Aplastic anemia, pancytopenia, thrombocytopenia, neutropenia, and leucopenia
7. Serious bleeding events
8. Lupus and lupus-like illness
9. Serious skin reactions (eg, Stevens Johnson Syndrome, toxic epidermal necrosis, and erythema multiforme)

Serious infections (1) will be summarized using the previously described serious TEAEs tables. No separate tables will be produced. In addition, summary tables will be produced for

opportunistic infections (including tuberculosis). Identification of opportunistic infections (including tuberculosis) will be performed using UCB-defined search criteria based on a two step process using identification via predefined list of PTs in addition to identification via manual review by the study physician.

(2) will be presented in two set of summary tables using the criteria standard MedDRA® query (SMQ) = "Malignant or unspecified tumours" and SMQ="Malignant tumours", respectively. The SMQ search will include all TEAEs which code to a PT included in the Scope=Narrow group within each SMQ. Corresponding summary tables will also include incidence of "Any malignancy (including unspecified, excluding non-melanomic skin cancers)" or "Any malignancy (excluding non-melanomic skin cancers)" depending on the table.

(3) will be presented in a stand-alone table and will be identified using the following UCB-defined search criteria:

- All serious TEAEs which code to a PT included in the search=Broad and or Narrow scope of the following SMQs:
  - Haemorrhagic central nervous system vascular conditions (SMQ)
  - Ischaemic central nervous system vascular conditions (SMQ) except events coding to PT "Transient ischaemic attack"
- All serious TEAEs which code to a PT included in the HLT "Ischaemic coronary artery disorders" except events coding to PT "Chest Pain" or "Chest discomfort"
- All serious TEAEs which code to a PT included in any of the following HLTs: "Heart Failures NEC", "Left Ventricular Failures", or "Right Ventricular Failures" and which also code to the SOC of "Cardiac Disorders" as Primary SOC.

(4) will be manually identified by the study physician from the previously described any TEAEs tables. No separate tables will be produced.

(5) will be presented in a stand-alone 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) will be presented in tables using the criteria SMQ = "Haematopoietic cytopenias" in the subset of Serious TEAEs. The SMQ search will include all serious TEAEs which code to a PT included in the Scope=Broad and/or Scope=Narrow groups within the SMQ.

(7) will be presented in tables using the criteria SMQ = "Haemorrhage terms (excl laboratory terms)" in the subset of serious TEAEs. The SMQ search will include all serious TEAEs which code to a PT included in the SMQ. Note that there is only a Narrow scope defined for this SMQ.

(8) will be manually identified by the study physician from the previously described any TEAEs tables. No separate tables will be produced.

(9) will be manually identified by the study physician from the previously described any serious TEAEs tables. No separate tables will be produced.

In addition, while not considered to be AEs of interest, following types of events will also be summarized:

- Hepatic events
- Hypersensitivity reactions and anaphylactic reactions

Hepatic events will be identified in the set of all TEAEs using following SMQs:

- Cholestasis and jaundice of hepatic origin (SMQ);
- Hepatic failure, fibrosis, and cirrhosis and other liver damage-related conditions (SMQ);
- Hepatitis, non-infectious (SMQ);
- Liver-related investigations, signs and symptoms (SMQ);
- Liver-related coagulation and bleeding disturbances (SMQ)

Hypersensitivity reactions and anaphylactic reactions will be summarized together. Corresponding set of tables will include incidence of following types of events:

- Any hypersensitivity and anaphylactic reactions
- Any hypersensitivity reactions
- Any anaphylactic reactions

All TEAEs that emerge within one day of when a study medication injection reaction was received which code to the following PTs will be considered to be a hypersensitivity reaction and included in the summary tables:

- 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
- Type IV hypersensitivity reaction

Injection reactions are identified by the investigator on the eCRF.

An algorithmic approach will be used to identify TEAEs that are considered to be anaphylactic reactions. PTs are separated into the 4 distinct categories (A, B, C, D) prior to the algorithmic approach being applied.

All TEAEs that emerge within one day of when a study medication injection reaction was received, and which fulfill any of the following 3 criteria will be included in the summary tables:

- If a subject reports any TEAE which codes to a PT included in Category A, then the event will be flagged as an anaphylactic reaction and summarized as such in the tables.

- If a subject reports any TEAE which codes to a PT included in Category B and reports any TEAE which codes to a PT included in Category C, and both TEAEs have the same start date, then both events will be flagged as anaphylactic reactions and summarized as such in the tables.
- If a subject reports any TEAE which codes to a PT included in Category D and reports (either a TEAE which codes to a PT included in Category B or a TEAE which codes to a PT included in Category C), and both TEAEs have the same start date, then both events will be flagged as anaphylactic reactions and summarized as such in the tables.

Category A consists of the following PTs which are included in the SMQ = Anaphylactic reaction (SMQ):

- Anaphylactic reaction
- Anaphylactic shock
- Anaphylactic transfusion reaction
- Anaphylactoid reaction
- Anaphylactoid shock
- Circulartory collapse
- Dialysis membrane reaction
- Kounis syndrome
- Shock
- Shock symptom
- Type I hypersensitivity

Category B consists of the following PTs:

- Acute respiratory failure
- Asthma
- Bronchial oedema
- Bronchospasm
- Cardio-respiratory distress
- Chest discomfort
- Chocking
- Chocking sensation
- Circumoral oedema
- Cough
- Cyanosis

- Dyspnoea
- Hyperventilation
- Irregular breathing
- Laryngeal dyspnea
- Laryngeal oedema
- Laryngospasm
- Lyngotracheal oedema
- Mouth swelling
- Nasal obstruction
- Oedema mouth
- Oropharyngeal spasm
- Oropharyngeal swelling
- Respiratory arrest
- Respiratory distress
- Respiratory dyskinesia
- Respiratory failure
- Reversible airways obstruction
- Sensation of foreign body
- Sneezing
- Stridor
- Swollen tongue
- Tachypnoea
- Throat tightness
- Tongue oedema
- Tracheal obstruction
- Tracheal oedema
- Upper airway obstruction
- Wheezing

Category C consists of the following PTs:

- Allergic oedema
- Angioedema

- Erythema
- Eye oedema
- Eye prutitus
- Eye swelling
- Eyelid oedema
- Face oedema
- Flushing
- Generalised erythema
- Injection site urticarial
- Lip oedema
- Lip swelling
- Nodular rash
- Ocular hyperaemia
- Oedema
- Periorbital oedema
- Pruritus
- Pruritus allergic
- Pruritus generalized
- Rash
- Rash erythematous
- Rash generalized
- Rash pruritic
- Skin swelling
- Swelling
- Swelling face
- Urticaria
- Urticaria papular

Category D consists of the following PTs:

- Blood pressure decreased
- Blood pressure diastolic decreased
- Blood pressure systolic decreased

- Cardiac arrest
- Cardio-respiratory arrest
- Cardiovascular insufficiency
- Diastolic hypotension
- Hypotension

The following AE summary tables (for the Parts/Analysis sets a, b, c, d) will be used as the basis for presentation of AEs of interest as well as for hepatic events, hypersensitivity reactions and anaphylactic reactions:

- All TEAEs by primary SOC, HLT, and PT (applies to opportunistic infections [1], [2], hepatic events, hypersensitivity reactions and anaphylactic reactions)
- All serious TEAEs by primary SOC, HLT, and PT (applies to [3], [5], [6] and [7])

EAIR and EAER will be calculated and included in these summary tables of AEs of interest.

Finally, a summary of all TEAEs which emerged at any time during the study (Part A or Part B) while on CZP 200mg Q2W will be produced. This summary will include all subjects in the SS and all subjects in the SSB who received CZP 200mg Q2W during Part B (either as randomized treatment or as escape treatment). EAIR and EAER will be calculated and included in this summary table.

TEAEs which emerged at any time during the study while on CZP 200mg Q2W will be defined as follows:

- In subjects not randomized into Part B, TEAEs under CZP full-dose therapy will be defined as all AEs starting on or after the date of first CZP full-dose therapy in Part A and up to 70 days after the last dose of study medication in Part A.
- In subjects randomized into Part B to the PBO or CZP 200mg Q4W arm who do not experience a flare and do not escape to CZP full-dose therapy, TEAEs under CZP full-dose therapy will be defined as all AEs starting on or after the date of first study medication in Part A and up to 70 days after the last dose of study medication in Part A and before the date of the first study medication in Part B.
- In subjects randomized into Part B to the PBO or CZP 200mg Q4W arm who do experience a flare and do escape to CZP full-dose therapy, TEAEs under CZP full-dose therapy will include all post-flare TEAEs in Part B.
- In subjects randomized into Part B to CZP 200mg Q2W arm, TEAEs under CZP full-dose therapy will include all TEAEs in Part A and Part B.

AE data will be listed. A glossary for the AE listing will detail the verbatim terms that are coded to each SOC, HLT, and PT.

10.3 Clinical laboratory evaluations

Testing for hepatitis B surface antigen and antibodies to hepatitis C and HIV will be performed at Screening. HLA-B27 will be performed at Screening only. These data will be listed only.

Descriptive statistics for observed values and change from Part A Baseline will be presented for each scheduled visit for the following parameters:

- Hematology: red blood cells, hemoglobin, hematocrit, platelets, white blood cells, neutrophils, lymphocytes, monocytes, eosinophils, and basophils
- Biochemistry: sodium, potassium, chloride, bicarbonate, total calcium, inorganic phosphorus, creatine phosphokinase (CPK), glucose, creatinine, urate, urea nitrogen, total protein, albumin, alkaline phosphatase (AP), aspartate aminotransferase (AST), alanine aminotransferase (ALT), bilirubin, total cholesterol

Values reported BLQ will be replaced by one-half the limit of quantification when reporting summary statistics.

Summaries for Part A will be presented by axSpA subpopulation and overall in the SS. Summaries for Part B will be presented by treatment group in the SSB excluding any post-flare assessments in the subset of subjects experiencing a flare in Part B and start escape therapy. For the subset of subjects experiencing a flare in Part B and start escape therapy, separate summaries will be produced using the ETS and will present only the post-flare assessments by treatment group.

For the continuous laboratory parameters, last value (end of treatment), minimum value during treatment, and maximum value during treatment, will be derived. Derivations will be performed for each treatment period separately. The derivations will only use post-baseline assessments in the corresponding treatment period and will exclude any assessment performed at the SFU visit. In Part B, derivations will be done excluding all assessments performed after the start of escape therapy in subjects experiencing a flare in Part B and start escape therapy. Additional derivations will be performed for the subset of subjects experiencing a flare in Part B and start escape therapy using the ETS and including only the assessments performed after the start of escape therapy.

The calculations for end of treatment, minimum and maximum values will be performed for each parameter separately.

A by-subject listing of all laboratory data will be provided. This listing will be presented by treatment group, results (with abnormal values flagged) and unit. In the listing for biochemistry also Glomerular Filtration Rate will be added.

All laboratory values, for which a normal range is available, will be classified into low (L) normal (N), and high (H) according to the respective normal range. Values falling outside of the normal limits will be classified as H or L and values within the range with N. In case no lower (upper) limit is given, the classification L (H) is not applicable.

A summary of shift from Baseline to the maximum, minimum and end of treatment value will be provided for hematology and biochemistry parameters for each treatment period separately (Part A, Part B pre-flare, Part B post-flare).

Presentations in Part A will be performed by axSpA subpopulation and overall in the SS. Presentations in Part B pre-flare will be performed by treatment group in the SSB. Presentations in Part B post-flare data will be performed by treatment group in the ETS.

Markedly abnormal values at each post-Baseline visit will be defined as laboratory values graded 3 or 4 according to the Rheumatology Common Toxicity Criteria (RCTC). Values fulfilling the criteria will be classified as marked abnormal high (MH) or marked abnormal low (ML). In case no lower (upper) limit is given the classification ML (MH) is not applicable.

Definitions of markedly abnormal values are given in the below Table 10.1 and Table 10.2 .

Table 10.1: Definitions of Markedly Abnormal Hematology Values

Parameter (SI units)	Markedly Abnormal Definition	
	Low	**High**
Hemoglobin (g/L)	<Lower Limit of Normal AND decrease from Baseline>20	N/A[a]
Hemoglobin (g/L)	<80	N/A
Leukocytes (total x 1000)	<2.0	N/A
Lymphocytes (x 1000)	<0.5	N/A
Neutrophils (x 1000)	<1.0	N/A
Platelets (x 1000)	<50	N/A

[a] N/A=not applicable.

Table 10.2: Definitions of Markedly Abnormal Biochemistry Values

Parameter (SI units)	Markedly Abnormal Definition	
	Low	**High**
AP	N/A[a]	>3 x ULN
ALT	N/A	>3 x ULN
AST	N/A	>3 x ULN
Calcium (mmol/L)	<1.75	>3.125
CPK	N/A	>4 x ULN
Creatinine	N/A	>1.8 x ULN
Glucose (mmol/L)	<2.22	>13.89
Potassium (mmol/L)	<3.0	>6.4
Sodium (mmol/L)	<125	N/A
Total bilirubin	N/A	≥2 x ULN
Uric acid	N/A	≥3 x ULN

[a] N/A=not applicable, ALT=alanine aminotransferase, AP=alkaline phosphatase, AST=aspartate aminotransferase, CPK=creatine phosphokinase.

The number and percentage of subjects with markedly abnormal hematology or biochemistry values will be summarized by visit and for any visit (excluding the SFU visit) separately for each treatment period. Subject numbers for subjects with any markedly abnormal hematology or biochemistry value will be tabulated.

Summaries for Part A will be presented by axSpA subpopulation and overall in the SS. Summaries for Part B will be presented by treatment group in the SSB excluding any post-flare assessments in the subset of subjects experiencing a flare in Part B and start escape therapy. For the subset of subjects experiencing a flare in Part B and start escape therapy, separate summaries will be produced using the ETS and will present only the post-flare assessments by treatment group.

For the derivation of the number and percentage of subjects with markedly abnormal hematology or biochemistry at any visit within a treatment period, only post-baseline assessments within the corresponding treatment period will be considered (excluding the SFU visit). In Part B, derivations will be done excluding all assessments performed after the start of escape therapy in subjects experiencing a flare in Part B and start escape therapy. In the subset of subjects experiencing a flare in Part B and start escape therapy, derivations will be done including only the assessments performed after the start of escape therapy.

A by-subject listing of markedly abnormal values will be provided separately.

The number and percentage of post-baseline liver function test evaluations will be summarized by varied ULN parameter criteria for ALT, AP, AST and total bilirubin separately for each treatment period, using same derivation as described previously. Presentations in Part A will be performed by axSpA subpopulation and overall using the SS. Presentations in Part B will be performed by treatment group using the SSB. For the subset of subjects experiencing a flare in Part B and do start escape therapy, presentations will be performed by treatment group using the ETS.

The varied parameter criteria will be:

- ≥2x, 3x, 5x, 10x, 20x ULN elevation of AST
- ≥2x, 3x, 5x, 10x, 20x ULN elevation of ALT
- ≥2x, 3x, 5x, 10x, 20x ULN elevation of AST or ALT
- ≥1x, 1.5x ULN elevation of Bilirubin
- ≥1.5x ULN elevation of AP
- ≥2x ULN elevation of Bilirubin and 3x ULN elevation of either ALT or AST

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

A by-subject listing for any subject with at least one post-baseline elevated test data will be provided. The listing includes all liver function test laboratory values for this subject.

Samples and results for additional laboratory tests (hepatitis B surface antigen and antibodies to hepatitis C and HIV, Genomics Laboratory Data, HLA-B27) will be listed separately.

The urinalysis data will be listed only.

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

10.4.1 Vital signs

Vital signs, including temperature, will be measured at all on-site visits (except at the 3 to 5 days prior to Week 48 Visit) including the SFU visit and also at all home-nurse visits in Part B. In subjects experiencing a flare in Part B, vital signs will be measured at escape Weeks 0, 2, 4, 12 and every 12 weeks thereafter. Respiration rate will be measured at Screening and Baseline only; and in addition at subsequent visits, if a subject experiences an AE.

Descriptive statistics for observed values and change from Part A Baseline will be presented for each scheduled visit for each vital sign parameter.

Vital signs will be presented for Part A and Part B separately. Part A summaries will be done by axSpA subpopulation and overall using the SS. Part B summaries will be done by treatment group using the SSB. Part B summaries will not include post-flare assessments. For the subset of subjects experiencing a flare in Part B and start escape therapy, summaries will be produced using the ETS including only the post-flare assessments. Vital signs assessments will be listed.

10.4.2 Other safety variables

The following presentations will be based on the SS for Part A and on the SSB for Part B, unless otherwise mentioned.

10.4.2.1 Pregnancy testing

Pregnancy testing must be carried out for women of childbearing potential and will consist of serum testing at Screening and SFU, and urine testing at Baseline and Week 96/withdrawal Visit. Pregnancy test results (i.e. Choriogonadotropin Beta) will be listed only.

10.4.2.2 Physical assessments

Physical examination will be performed at Screening, Baseline, Weeks 12, 24, 36, 48, 60, 72, 84, 96/withdrawal, and at the SFU visit (10 weeks after the last dose). In subjects experiencing a flare in Part B, physical examination will be performed at escape Weeks 0, 2, 4, 12 and every 12 weeks thereafter. Clinically important abnormal changes in subsequent physical examinations will be recorded as AEs. Physical examination results will be listed.

Weight is to be measured at Screening, Baseline, Week 48, and at completion at Week 96/withdrawal Visit. Height will be measured at the Baseline Visit only. Summary statistics will be provided by axSpA subpopulation and overall in Part A using the SS and by treatment group in Part B using the SSB by visit presenting actual values and changes from Part A Baseline for weight, height and BMI, as applicable. The summary for Part B will not include post-flare assessments. For the subset of subjects experiencing a flare in Part B and start escape therapy, a summary by treatment group will be produced using the ETS including the post-flare assessments. These data will also be listed.

10.4.2.3 Tuberculosis assessments

During the conduct of the study, the tuberculosis (TB) assessment by IGRA (QuantiFERON TB GOLD In Tube test or Elispot, if QuantiFERON TB GOLD In Tube test is not available locally), will be repeated at Week 48 and Week 96/withdrawal Visit for all subjects.

Tuberculosis testing results will be listed.

A plain posteroanterior chest x-ray (or, if done, computed axial tomography of the chest) must be done within 3 months prior to the Screening Visit. The chest x-ray should be repeated only if the TB test was confirmed positive or any further evidence is suggestive of potential TB infection (eg, exposure).

Chest x-ray results will be listed.

11 REFERENCES

Guyatt G, Mitchell A, Irvine E J, Singer J, Williams N, Goodacre R, Tompkins C. A new measure of health status for clinical trials in inflammatory bowel disease. Gastroenterology. 1989 Mar; 96(3) 804-10.

Heinze, G. and Schemper, M. A Solution to the Problem of Separation in Logistic Regression. Statistics in Medicine, 2002; 21, 2409–2419.

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; 62:127-32.

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.

Mallinckrodt C.H. Preventing and Treating Missing Data in Longitudinal Clinical Trials. A Practical Guide. Cambridge University Press, 2013.

Siddiqui O, Hung HMJ, O'Neill R. MMRM vs. LOCF: A comprehensive comparison based on simulation study and 25 NDA datasets. J Biopharm Stat. 2009; 19: 227-246.

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

12 APPENDICES

12.1 Rules for medical resource utilization

12.1.1 Healthcare Resource Utilization

Study-specific modules will be used to capture the following data regarding healthcare resource utilization during the study:

- Concomitant medical procedures
- Health care provider consultations not foreseen by the protocol
- In-patient hospitalizations and emergency room visits (including length of stay)

In addition to the above categories, the duration of hospital stay will be calculated.

12.1.2 Counting rules

12.1.2.1 In-patient hospitalization and emergency room visits

1. In case of complete admission and discharge dates, for the same subject only one hospitalization will be considered if "start date of the second hospitalization - end date of the first hospitalization≤1".
2. In case the discharge date for a hospitalization is missing (regardless of the discharge check value),

- if there is a subsequent hospitalization,
  - if the initial entry point and relationship are the same between the 2 hospitalizations, then the first hospitalization will be grouped with the next one

    (1) if the discharge date of the second hospitalization is non-missing, then length = second hospitalization discharge date – first hospitalization admission date +1;

    (2) else,

    length = last non missing visit date – first hospitalization admission date +1 .
  - otherwise, if either the entry point or the relationship of the 2 hospitalizations is different, then both hospitalizations will be counted as distinct. In this case the length of the first hospitalization will be computed as follows:

    length = (second hospitalization admission date – 1) – first hospitalization admission date + 1= second hospitalization admission date – first hospitalization admission date.
- in case there is no subsequent hospitalization, then the length of the hospitalization is calculated as

  length = last non missing visit date – hospitalization admission date +1.

3. In case the last visit date is before the start date, then the discharge date for a hospitalization is replaced by the start date of hospitalization and the length of the hospitalization is equal to 1.

12.1.2.2 Healthcare provider consultations not foreseen by the protocol

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

12.1.2.3 Concomitant medical procedures

If the procedure name, start date, and relationship are the same, then only 1 procedure is counted, otherwise if at least 1 variable among procedure name, start date, or relationship is different, distinct/several procedures are counted.

12.2 Compliance ratio calculation

The general formula for treatment compliance is given as follows:

Compliance Ratio (CR)= {[Exposure Duration (days)] - [Cumulative Difference (days)]}/

{Exposure Duration (days)}

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

CR(1):

Exposure Duration (days) = Last Visit date – Baseline date + 14

Cumulative Difference (days)=$\sum_i$ ABS($AD_i - SD_i$), where i=Week 0, 2, 4, 6, … , XX

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

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

CR(2):

Exposure Duration (days) = Last Visit date – Baseline date +14

Cumulative Difference (days)=$\sum_i$ ABS($AD_i - SD_i$), where i= Week 0, 2, 4, 6, … , XX

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

$\{14 + AD_{i-2},$ for Week i=2 to XX by 2 $\}$

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

Table 12.1: Example

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

Note that this example is for a 12 week study.

In case the actual day of administration is missing, the maximum deviation to the scheduled visit date will be assumed (ie 14 days). If for a scheduled visit date with 2 planned injections, the syringes were administered on 2 different days, the maximum day difference of the 2 actual dates to the scheduled visit date will be utilized.

For the measurement of treatment compliance in Part A, the exposure duration in Part A, identified as “Exposure Duration [a]”, will use the last of the open label injection visit dates:

- Exposure Duration [a] (days) = week 46 Visit/last injection date – Baseline date + 14

The cumulative difference will be calculated for all visits from Week 0 (Baseline) to week 46.

For the measurement of treatment compliance in Part B, the exposure duration in Part B, identified as “Exposure Duration [b]”, will use the last of the double-blind injection visit dates:

- Exposure Duration [b] (days) = week 94 Visit/last injection date – Week 48 injection/Week 48 Visit date + 14

The cumulative difference will be calculated for all visits from Week 48 to week 94.

For the measurement of treatment compliance in Part B post-flare, the post-flare exposure duration, identified as “Exposure Duration [f]”, will use the last of the injection escape Visit dates:

- Exposure Duration [f] (days) = last escape visit/last injection date – first Escape Week 0 injection /Escape Week 0 Visit date + 14

The cumulative Difference will be calculated for all visits from Escape Week 0 to last Escape visit.

12.3 SAP Amendment 1

Rationale for the amendment

The main purpose of this amendment is to implement the updates of Protocol Amendment 2 where the current assay for measuring ADAb in plasma was replaced by a new method in order to align with current regulatory guidelines.

In line with the Protocol of the study, the Safety Set changed to include the subjects in open-label Part A as well. The previous defined Safety Set was changed to Safety Set Part B (SSB).

There is also the possibility to report one unscheduled ASDAS between two scheduled ASDAS assessments, the censoring time point changed and additional instruction for analysis are added.

Moreover, scores which should present an improvement (ASDAS-CII, ASDAS-MI, ASAS20, ASAS40, ASAS5/6, SF-36 PCS response and SF-36 MCS response) are not analyzed using Part B Baseline as reference.

A modified version of ASDAS-MI was added as sensitivity analysis for ASDAS-MI.

Analyses for time since diagnosis of disease and symptom duration were added to medical history.

Modifications and changes

Global changes

The following changes were made throughout the SAP:

- In the summary tables for Part A, the axSpA subpopulation is used instead of mNY classification, since the data is presented for AS and nr-axSPA in the corresponding outputs. The term mNY classification is kept when it is presented or used as stratification factor.
- In safety outputs for Part A, SS is used instead of OLS and in safety outputs for Part B SSB is used instead of SS.
- Censoring point changed from week 92 to week 94.
- Analyses for ASDAS-CII, ASDAS-MI, ASAS20, ASAS40, ASAS5/6, SF-36 PCS response and SF-36 MCS response referring to Part B Baseline are removed.

Specific changes

In addition to the global changes, the following specific changes have been made (formats as missing spaces or redundant spaces are not listed):

Change #1

List of abbreviation

The following abbreviations were added:

Nab Neutralizing antibody

SSB Safety Set Part B

ADAb anti-CZP antibody

Change #2

1 Introduction

The following protocol version was added:

- Protocol Amendment 2 dated 24 Jan 2018
- Protocol Amendment 2.1 (Taiwan) dated 27 Feb 18
- Protocol Amendment 2.2 (UK) dated 27 Feb 18

Change #3

2.2.1.3.2 Other efficacy variables for subjects entering Part B

- ASDAS clinical improvement (ASDAS-CII, ASDAS-MI) using Part A Baseline and Part B Baseline as a reference
- ASAS20, ASAS40, ASAS5/6, and ASAS PR response using Part A Baseline Baseline and Part B Baseline as a reference
- BASDAI50 response using Part A Baseline and Part B Baseline as a reference
- SF-36 PCS response and SF-36 response using Part A Baseline and Part B Baseline as a reference

Has been changed to:

- ASDAS clinical improvement (ASDAS-CII, ASDAS-MI) using Part A Baseline as a reference
- ASAS20, ASAS40, ASAS5/6, and ASAS PR response using Part A Baseline as a reference
- BASDAI50 response using Part A Baseline as a reference
- SF-36 PCS response and SF-36 MCS response using Part A Baseline as a reference

Change #4

2.2.1.3.3 Other efficacy variables for subjects who experience a flare in Part B

- ASDAS clinical improvement (ASDAS-CII, ASDAS-MI) using Part B Baseline and Flare Baseline as reference
- ASAS20, ASAS40, ASAS5/6, and ASAS PR response using Part B Baseline and Flare Baseline as reference
- BASDAI50 response using Part B Baseline and Flare Baseline as reference
- SF-36 PCS response and SF-36 response using Part B Baseline and Flare Baseline as a reference

Has been changed to:

- ASDAS clinical improvement (ASDAS-CII, ASDAS-MI) using Flare Baseline as reference
- ASAS20, ASAS40, ASAS5/6, and ASAS PR response using Flare Baseline as reference

- BASDAI50 response using Flare Baseline as reference
- SF-36 PCS response and SF-36 MCS response using Flare Baseline as a reference

Change #5

2.2.2.2 Immunogenicity variables

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

- Number and % of subjects with Ab >2.4 units/mL at the time of each visit
- Number and % of subjects with Ab >2.4 units/mL at any visit during treatment (not including the safety follow-up (SFU) visit)
- Number and % of subjects with Ab >2.4 units/mL at any visit including the SFU visit

For the subgroup of subjects with at least 1 anti-CZP Ab >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) levels will be assessed at Baseline and subsequent time points as described in the protocol.

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

Change #6

3.2.4 MRI assessments

MRI will be assessed centrally and scoring will be done by 2 independent readers, who are blinded to both the order of the scans and to the treatment group, using a previously reported scoring system. Two different MRI scores will be used for analysis.

Has been changed to:

Magnetic resonance imaging of the spine and sacroiliac joints will be performed at Screening, Week 48, Week 96, or WD Visit if MRI was performed more than 12 weeks prior to WD. MRI will be assessed centrally and scoring will be done by 2 independent readers, who are blinded to both the order of the scans and to the treatment group, using a previously reported scoring system. In addition, a single reader assessment will be performed for change from Baseline (Week 48) evaluation in the MRI scores for all subjects that entered Part A. Two different MRI scores will be used for analysis.

Change #7

3.2.7 Ankylosing Spondylitis Disease Activity Score (ASDAS)

0.579 × (natural logarithm of the CRP [mg/L] + 1)

Has been changed to:

0.579 × (natural logarithm of the (CRP [mg/L] + 1))

Change #8

3.2.7 Ankylosing Spondylitis Disease Activity Score (ASDAS)

- ASDAS-MI: ASDAS reduction (improvement) of ≥2.0 relative to Baseline.

Has been changed to:

- ASDAS-MI: ASDAS reduction (improvement) of ≥2.0 relative to Baseline. As a sensitivity analysis, ASDAS-MI analyses will be repeated using a Modified ASDAS-MI: ASDAS reduction (improvement) of >=2.0 relative to Baseline or has the lowest score possible post-baseline (i.e. when CRP<LLOQ and all other components are 0, then the minimum ASDAS score is 0.636 to 3 decimal places).

Change #9

3.2.22 Flare

The following part has been added:

(in case of an unscheduled ASDAS visit at Week 94).

Change #10

3.2.22 Flare

The following part has been added:

There is also the possibility to report one unscheduled ASDAS between two scheduled ASDAS assessments. The same rules as for the scheduled ASDAS apply for the unscheduled ASDAS.

The unscheduled ASDAS will be taken into account for flare calculation for two consecutive visits if there are more than 13 calendar days between two consecutive ASDAS calculated dates.

If the ASDAS score is missing at 2 scheduled visits and unscheduled ASDAS entry present in between and not larger than 3.5, then the subject will not be considered to have experienced a flare.

Change #11

3.2.23 Time to flare

The following part has been added:

- If the assessments were not performed on the same day, a conservative approach will be used and the latest day will be taken as calculated ASDAS day.

Change #12

3.2.23 Time to flare

- For subjects who complete the study without meeting the criteria for flare, the censored date will be the date of Week 12 visit.

Has been changed to:

- For subjects who complete the study without meeting the criteria for flare, the censored date will be the last possible date of the event, i.e. Week 92 for scheduled ASDAS and Week 94 if an unscheduled ASDAS was performed at Week 94.

Change #13

3.2.26 Anti-CZP antibody status

The anti-CZP Ab status is positive (Ab+) at a given visit if the Ab level is >2.4 units/mL. The anti-CZP Ab status is negative (Ab-) at a given visit if the Ab level is ≤2.4 units/mL.

A subject has an overall positive anti-CZP Ab status if the Ab level is >2.4 units/mL on at least 1 visit during treatment, excluding the SFU Visit. A subject has an overall negative anti-CZP Ab status if the Ab level is ≤2.4 units/mL at all visits during treatment, excluding the SFU visit.

Has been changed to:

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

Change #14

3.2.27 Modified New York criteria

The following part has been added:

The mNY classification will determined at Screening. MRI will be assessed centrally and by 2 independent readers. If the results from the readers are different, then the result of the adjudicator will be used.

Change #15

3.3 Analysis time points

The following parts have been added:

If a subject does not have a Week 96/withdrawal Visit, then safety and efficacy data recorded at either the date of the last scheduled or unscheduled visit during the treatment Period, whichever is later, will be included in summaries “Last/Withdrawal” visit.

…

If a subject does not have a Withdrawal Visit, then safety and efficacy data recorded at either the date of the last scheduled or unscheduled visit during the Escape Period or the date of last known dose during the Escape Period, whichever is later, will be included in summaries “Last/Withdrawal” visit.

Change #16

3.6 Analysis sets

There will be 10 analysis sets used within this study

Has been changed to:

There will be 11 analysis sets used within this study

Change #17

3.6.4 Safety Set

The Safety Set (SS) will consist of all subjects in the RS who have received at least 1 dose of study medication in the Double-Blind Period of the study (Part B).

Has been changed to:

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

Change #18

The following section has been added:

3.6.5 Safety Set Part B

The Safety Set Part B (SSB) will consist of all subjects in the RS who have received at least 1 dose of study medication in the Double-Blind Period of the study (Part B).

Change #19

3.6.7 Pharmacokinetic Set A

The Pharmacokinetic Set A (PKSA) will consist of all subjects from the OLS who provide at least 1 PK sample during Part A.

Has been changed to:

3.6.8 Pharmacokinetic Set A

The Pharmacokinetic Set A (PKSA) will consist of all subjects from the SS who provide at least 1 PK sample during Part A.

Change #20

3.6.8 Pharmacokinetic Set B

The Pharmacokinetic Set B (PKSB) will consist of all subjects from the SS who provide at least 1 PK sample during Part B.

Has been changed to:

3.6.9 Pharmacokinetic Set B

The Pharmacokinetic Set B (PKSB) will consist of all subjects from the SSB who provide at least 1 PK sample during Part B.

Change #21

The following section was added:

3.6.13 General use of the Analysis Sets

Efficacy summaries for variables collected in Part A will be based on the OLS. All efficacy analyses for Part B will be performed using the RS. The FAS and PPS will be used for a sensitivity analysis on the primary endpoint only. The other efficacy variables for subjects who experience a flare in Part B will be based on the FS.

All safety summaries in Part A will be based on the SS. All safety summaries in Part B will be based on the SSB. The other safety variables for subjects who experience a flare in Part B will be based on the ETS.

Pharmacokinetic summaries will be based on to Part A/ Part B corresponding set.

Change #22

3.8 Center pooling strategy:

Italy was removed in the list of countries in Table 3-3.

Change #23

3.10 Changes to protocol-defined analyses

The protocol defines the SS as all subjects in the ES who have received at least 1 dose of study medication. The definition of the SS in the SAP has changed from that in the protocol. According to the SAP, the SS only includes subjects in the RS who have received at least 1 dose of study medication in the Double-Blind Period of the study (Part B). The rational for the modification of this definition is that using the protocol definition, the SS would be identical to the OLS.

Has been changed to:

The protocol defines the censor date for time flare as week 96. The censor date in the SAP is defined as the last possible visit of event, i.e. week 94.

In addition to the protocol, a modified version for ASDAS-MI is added.

Change #24

4.2 Handling of dropouts or missing data

The following part has been added:

Specifically, if a subject withdraws from Part B at any point, they will be considered as having not responded to treatment. Therefore, they will be assumed to have experienced a flare during Part B for analysis purposes.

Change #25

4.2 Handling of dropouts or missing data

For all analyses of binary endpoints (sustained remission, ASDAS-MI, ASDAS-CII, ASAS20, ASAS40, ASAS5/6, ASAS PR, BASDAI50 response, IBD-Q remission, IBD-Q response, SF-36 PCS response, SF-36 MCS response) NRI will be used. Thus, unless mentioned otherwise (see Section 3.2.2 for ASAS20, ASAS40, ASAS5/6 and ASAS PR), a subject having missing data for the time point assessed will be conservatively counted as a non-responder. This will be done whether the data is missing, the subject discontinued prior to the time point assessed, or the data is considered missing due to start of escape treatment.

Has been changed to:

For all analyses of binary endpoints (sustained remission, ASDAS-MI, ASDAS-CII, modified ASDAS-MI, ASAS20, ASAS40, ASAS5/6, ASAS PR, BASDAI50 response, IBD-Q remission,

IBD-Q response, SF-36 PCS response, SF-36 MCS response) NRI will be used. Thus, unless mentioned otherwise (see Section 3.2.2 for ASAS20, ASAS40, ASAS5/6 and ASAS PR), a subject having missing data for the time point assessed will be conservatively counted as a non-responder. This will be done whether the data is missing, the subject discontinued prior to the time point assessed, or the data is considered missing due to start of escape treatment.

Change #26

4.2 Handling of dropouts or missing data

- For missing post Week 48 Part B assessments, only carry forward earlier Part B post-Week 48 values from assessments performed before the start of escape therapy. Week 48 values do not get carried forward to the Part B post-baseline visits

Has been changed to:

- For missing post Week 48 Part B assessments,
  - if the subject did not experience a flare in Part B, only carry forward earlier Part B post-Week 48 values,
  - if the subject did experience a flare in Part B, the assessment at escape week 0 will be treated as the last assessment in Part B prior to escape treatment. For any missing post 48 Part B assessments up to the week of flare (the later visit if flare was based on 2 visits), carry forward earlier Part B post-Week 48 values. Assessments whilst on escape treatment are not to be used, therefore the assessment at escape week 0 will be carried forward for all subsequent visits until Week 96. E.g if the subject experienced a flare at week 72 then the Escape week 0 assessment will be carried forward for the assessments after week 72 in Part B (76, 80, 84, 88, 92 and 96)

  Week 48 values do not get carried forward to the Part B post-baseline visits.

- Assessments during escape treatment will not be carried forward.

Change #27

A database lock may be performed and an interim study report may be written, if required. The scope of an Interim Analysis will be covered in a separate Analysis Plan.

Has been changed to:

4.3 Interim analyses and data monitoring

The purpose of the interim analysis is to check whether the assumptions on efficacy of CZP made for the protocol were correct, and to potentially present the results at scientific meetings. Main focus of this interim analysis will be the efficacy and safety up to Week 48. An interim freeze of the database will be performed once all subjects have completed Week 48 of Part A. Part A data will be reported in corresponding Part A tables and listings. All blinded data will remain blinded. In the disposition tables, treatment will remain blinded and be presented according to a dummy treatment assignment for Part B. For the listing only open-label data from Part A will be included. No CSR will be written. The following assessments will not be included in the Interim Analysis:

- MRI Efficacy scores (SPARCC (Sacroiliac joint), Spine ASspiMRI-a in Berlin Modification score)
- Work Productivity Questionnaire
- EQ-5D-3L
- Fecal and serum calprotectin
- Correlation between calprotectin and IBD, CRP
- SF36 (domains, PCS, MCS)
- Resource utilization
- PK
- ADAb

The scope of the Interim Analysis will not be covered in a separate Analysis Plan.

Change #28

4.7 Examination of subgroups

- Overall anti-CZP Ab status (Ab-, Ab+)

Has been changed to:

- ADAb status

Change #29

5.2 Protocol deviations

blinded data review meeting

Has been changed to:

blinded data evaluation meeting

Change #30

The following part has been added:

Time since diagnosis of disease and symptom duration will be summarized by axSpA subpopulation and for all subjects combined in Part A using the SS, by treatment group and for all subjects combined in Part B using the SSB. Time since diagnosis will also be categorized into ≤2 and >2 years, and symptom duration will be categorized into ≤3 and >3 years.

Time since diagnosis of disease will be defined as: Earliest start date of the medical history of any of the preferred terms ‘Axial Spondyloarthritis’, 'Spondyloarthropathy' , 'Ankylosing spondylitis' , or 'Spondylitis' minus date of first study medication administration. If a subject does not have a history of any of the above 4 preferred terms, then the time since diagnosis of disease will be set equal to the symptom duration.

Symptom duration will be calculated as:

Date of first study medication administration – Start date of symptoms

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 for concomitant medications in Section 4.2):

- Subjects with a medical history preferred term of 'Back pain', 'Inflammatory pain', 'Spinal Pain', 'Lumbar spine 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', 'Spondyloarthropathy', 'Ankylosing spondylitis', or 'Spondylitis' 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.

If the resulting symptom duration is unknown or is below 3 months, the imputed start date will be set to 3 months and 1 day prior to the start of treatment so that the symptom duration is imputed as 3 months.

Change #31

8.3.2.9 Number of uveitis flares, IBD exacerbations and psoriasis exacerbations

For each of these events (uveitis flares, IBD exacerbations and psoriasis exacerbations), the summary table will be repeated based on the subset of subjects with a current occurrence of the event at screening.

Has been changed to:

For each of these events (uveitis flares, IBD exacerbations and psoriasis exacerbations), the summary table will be repeated based on the subset of subjects with a history of occurrence and/or a current occurrence of the event at screening.

Change #32

9.2 Immunogenicity

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

- Number and % of subjects with Ab >2.4 units/mL at the time of each visit
- Number and % of subjects with Ab >2.4 units/mL at any visit during treatment (not including the SFU visit)
- Number and % of subjects with Ab >2.4 units/mL at any visit including the SFU visit

Summaries will be done separately in Part A by mNY and overall using PKSA and in Part B by treatment group using PKSB. Part B summary will not include post-flare assessments. Separate summary by treatment group will be produced for subjects who experience a flare in Part B and start escape therapy including only the post-flare assessments.

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

Has been changed to:

Immunogenicity will be assessed through listing of individual results by subject and summary tables. Immunogenicity data will be correlated with PK and efficacy readout. In addition, immunogenicity will be correlated with possible safety findings.

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 ADAb assay..

Subjects will receive an overall classification, inclusive of Baseline and Post-Baseline results, and be classified as follows based on the ADAb assay results:.

- ADAb negative: no confirmed positive ADAb samples at any of the sampling time points.

ADAb positive: confirmed positive ADAb samples at one or more sampling time points.

Summaries will be done separately in Part A by axSpA subpopulation and overall using PKSA and in Part B by treatment group using PKSB. Part B summary will not include post-flare assessments. Separate summary by treatment group will be produced for subjects who experience a flare in Part B and start escape therapy including only the post-flare assessments.

Change #33

10.2 Adverse events

(5) will be manually identified by the study physician from the previously described any TEAEs tables. No separate tables will be produced.

Has been changed to:

(5) will be presented in a stand-alone 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.

Change #34

10.2 Adverse events

- All serious TEAEs by primary SOC, HLT, and PT (applies to [3], [6] and [7])

Has been changed to:

- All serious TEAEs by primary SOC, HLT, and PT (applies to [3], [5], [6] and [7])

Change #35

12.2.2.1 In-patient hospitalization and emergency room visits

The following part has been added:

- In case the last visit date is before the start date, then the discharge date for a hospitalization is replaced by the start date of hospitalization and the length of the hospitalization is equal to 1.

12.4 SAP Amendment 2

Rationale for the amendment

The main purpose of this amendment is to specify the ADAb analysis. The CZP plasma concentration and ADAb titer of the patients that did not reach sustained remission are not analyzed, the PKSA as described in the protocol will be removed from the SAP.

In line with the charter with medical imaging for MRI reads the Section 3.2.4 was updated.

Further specification for the analysis of flare and time to flare are added (Subject with incorrectly identified flared in Part B, seeds for referenced based multiple imputation of ASDAS).

The "Last/Withdrawal" visit and the "Last" visit for Escape Part were renamed to "Last/Withdrawal Escape" visit and "Last Visit (Week 96 Termination)" for clarification and to indicated that summary statistics consist of different data.

Definition of Part B Baseline changed and further specification in all baseline definitions were added.

Scoring software for SF-36 was updated calculate all subjects with the same software.

The model for MMRM was specified for changes to Part A and Part B Baseline.

Imputation rules for medical history were specified

Modifications and changes

Global changes

The following changes were made throughout the SAP:

- Censoring point changed from week 94 to week 96 in accordance with the protocol
- ASDAS-Moderate Disease changed to ASDAS-Low Disease activity (Machado et al. 2018).
- LOCF analysis are removed from Part B efficacy tables and for selected Part A efficacy tables.
- MMRM will be used instead of ANCOVA based on LOCF for
  - Total spinal pain
  - Nocturnal spinal pain
  - PtGADA

  - Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
  - SF-36 PCS, MCS, and physical function domain
  - Fatigue NRS
  - ASQoL
- Exploratory statistical comparisons for WPS will be based on MMRM instead of nonparametric bootstrap-t method.
- List of efficacy variables in section 2.2.1 are aligned with order in the Protocol and the analysis for IBD-Q response and IBD-Q remission using Part A Baseline and Part B Baseline as a reference and SF-36 PCS response and SF-36 MCS response are removed from the list.
- In addition, for the primary variables and selected outputs the 95% confidence interval (CI) based on the Exact Binomial distribution will be added.
- Wording from "subject who flare" changed to "subjects who experiencing a flare".
- Subsections in Section 3.2.22 added to provide a better overview about the different imputation methods for sensitivity analysis of the primary endpoint.
- Selected baseline characteristic tables based on the Open-Label Set will be repeated for the Randomized Set

Specific changes

In addition to the global changes, the following specific changes have been made (formats as missing spaces or redundant spaces are not listed, typos):

Change #1

List of abbreviation

ASAS	Assessment of SpondyloArthritis International Society
ASAS20, 40 criteria	Assessment of SpondyloArthritis International Society 20%, 40% response
AS	Ankylosing Spondylitis
ASDAS-MD	Ankylosing Spondylitis Disease Activity Score – Moderate Disease
ASspIMRI-a	Ankylosing Spondylitis spine MRI score for activity
ATC	Anatomical Therapeutic Chemical
CPK	Creatine Phosphokinase
DMARD	Disease-modifying antirheumatic drug
EAIR	Exposure adjusted incidence rate
EAER	Exposure adjusted event rate
IXRS	Interactive voice or web response system
NRS	Numeric Rating Scale

WHO-DRL	World Health Organization Drug Reference List

Has been changed to:

ASAS	Assessment of SpondyloArthritis international Society
ASAS20, 40 criteria	Assessment of SpondyloArthritis international Society 20%, 40% response
AS	ankylosing spondylitis
ASDAS-LD	Ankylosing Spondylitis Disease Activity Score – Low Disease activity
ASspiMRI-a	Ankylosing Spondylitis spine MRI score for activity
ATC	anatomical therapeutic chemical
CPK	creatine phosphokinase
DMARD	disease-modifying antirheumatic drug
EAIR	exposure adjusted incidence rate
EAER	exposure adjusted event rate
IXRS	interactive voice or web response system
NRS	numeric rating scale
WHO-DD	World Health Organization Drug Dictionary

The following abbreviation has been added:

ANCOVA	analysis of covariance
BASDAI50	Bath Ankylosing Spondylitis Disease Activity Index 50% response criteria
CI	confidence intervaleCRF electronic Case Report Form
FS	Flared Set
H	high
HIV	human immunodeficiency virus
HLA-B27	human leukocyte antigen B27
HLT	High Level Term
IGRA	interferon-gamma release assay
L	low
MAR	missing at random
MSR	minimum significant ratio
N	normal
NA	not applicable
PK	pharmacokinetics

PBO	placebo
PR	partial remission
PT	preferred term
Q2W	every 2 weeks
Q4W	every 4 weeks
r-axSpA	radiographic axial spondyloarthritis
SOC	system organ class
STIR	short-tau-inversion recovery
TB	tuberculosis
TE	treatment-emergent
UK	United Kingdom
VU	vertebral units

The following abbreviation has been removed:

Ab	antibody
LLN	lower limit of normal

Change #2

2.2.2 Pharmacokinetic immunogenicity variables

2.2.2.1 Pharmacokinetic variables

Plasma concentrations of CZP will be measured at Baseline and subsequent time points as described in Table 5.1 and Table 5.2 of the protocol. These plasma samples may be used additionally for analyses of CZP and its constituent moieties using alternative methods.

2.2.2.2 Immunogenicity variables

Anti-CZP antibody (ADAb) levels will be assessed at Baseline and subsequent time points as described in the protocol.

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

Has been changed to:

2.2.2 Pharmacokinetic, pharmacogenomic and immunogenicity variables

2.2.2.1 Pharmacokinetic variables

Plasma concentrations of CZP will be measured at Part A Baseline and subsequent time points as described in Table 5.1 and Table 5.2 of the protocol only for subjects randomized into Part B (see Section 3.10). These plasma samples may be used additionally for analyses of CZP and its constituent moieties using alternative methods.

2.2.2.2 Biomarkers

Any analysis for the biomarker will be described in a separate document.

2.2.2.3 Pharmacogenomic variables

Analysis for the pharmacogenomic will be described in a separate document.

2.2.2.4 Immunogenicity variables

Anti-CZP antibody (ADAb) levels will be assessed at Part A Baseline and subsequent time points as described in the protocol.

Determination of ADAb will be done using a validated screening, confirmation, and titration ADAb bridging assay, with potential further characterization by a neutralizing antibody (NAb) assay. The NAb will be not part of the CSR. The immunogenicity samples will be processed according to dedicated Bioanalytical Analysis plans.

Change #3

3.1 General presentation of summaries and analyses

The following parts have been added:

In addition, for the primary variables and selected outputs the 95% confidence interval (CI) based on the Exact Binomial distribution will be presented.

…

The study design has three different periods: Part A, Part B and Escape Part. The outputs will be presented for each period. Summaries for Part A will be presented by axSpA subpopulation. Summaries for Part B will be presented by treatment group excluding any post-flare assessments in the subset of subjects who experiencing a flare in Part B and start escape therapy. For the subset of subjects who experiencing a flare in Part B and start escape therapy, separate summaries will be produced for the Escape Part.

Change #4

3.2.2 ASAS20, ASAS40, ASAS 5/6 response, and ASAS Partial Remission

The following part has been added:

The lateral lumbar flexion value is taken as the mean of left and right assessment values and improvement is a ≥20% increase. Improvement in CRP is a ≥20% decrease.

Change #5

3.2.4 MRI assessments

Magnetic resonance imaging of the spine and sacroiliac joints will be performed at Screening, Week 48, Week 96, or WD Visit if MRI was performed more than 12 weeks prior to WD. MRI will be assessed centrally and scoring will be done by 2 independent readers, who are blinded to both the order of the scans and to the treatment group, using a previously reported scoring system. In addition, a single reader assessment will be performed for change from Baseline (Week 48) evaluation in the MRI scores for all subjects that entered Part A. Two different MRI scores will be used for analysis.

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 SIJ 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 are 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 SIJ 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.

…

In case of out of window images/scores regarding the visit schedule, the following rules will apply:

Has been changed to:

Magnetic resonance imaging of the spine and sacroiliac joints will be performed at Screening, Week 48, Week 96, or WD Visit if the previous MRI was performed more than 12 weeks prior to WD. In the Escape Part it will be also performed at Escape Week 0 and 12. MRI will be assessed centrally and scoring will be done according to the imaging charter by 2 independent readers, who are blinded to both the order of the scans and to the treatment group, using a previously reported scoring system. A single reader assessment will be performed for change from Baseline (Week 48) evaluation in the MRI scores for all subjects that entered Part A. A double reader assessment will be performed for subjects entering Part B. The analysis will use the average of the scores from the 2 independent reviewers. If the two readers disagree a third reviewer who is not one of the two reviewers will perform adjudication by providing a third independent review. Whenever an adjudication, is present the average score across all 3 reviewers will be used for analysis. In addition, a single reader assessment will be performed for change from Baseline (Week 48) evaluation in the MRI scores for all subjects that entered Part A. Two different MRI scores will be used for analysis (SPARCC, Berlin modification of the ASspiMRI-a).

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 SIJ is divided into 4 quadrants and will be assessed for left and right: upper iliac, lower iliac, upper sacral, and lower sacral. The presence of increased signal on STIR in each of these 4 quadrants are 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 SIJ 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) or 24 VUs if transitional vertebra is present. If a transitional vertebra is present L5/L6 and L6/SI will be assessed instead L5/SI. 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. This score will increase if there are transitional vertebrae.

…

In case of out of window images/scores regarding the visit date, the following rules will apply:

Change #6

3.2.7 Ankylosing Spondylitis Disease Activity Score (ASDAS)

The following part has been added:

For Part A Baseline derivation of the ASDAS score the latest value prior to treatment will be used for each component. It will be not required that all components are measured at the same visit, e.g. PtGADA is available only at screening while the other components are available at Baseline then the ASDAS score will be derived combining the screening value of PtGADA and the baseline value of the other components.

For Part B Baseline derivation of the ASDAS score the latest value prior to randomization will be used for each component. It will be not required that all components are measured at the same visit.

For Flare Baseline derivation of the ASDAS score the latest value prior to first injection in Escape Part will be used for each component. It will be not required that all components are measured at the same visit.

Change #7

3.2.11 Short-Form 36-item Health Survey (SF-36)

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

Has been changed to:

For the calculation of the SF-36 domain scores and the component summaries PCS and MCS, the scoring software Optum's PRO CoRE will be used. The norm-based scores (based on the US 2009 general population) will be utilized for analysis.

Change #8

3.2.22 Flare

The following part has been added:

If a subject was incorrectly identified as flared in Part B at any point in Part B due the to late data entry in IXRS aand went to escape Part B, then the subject will be considered to have experienced a flare.

Change #9

Multiple imputation of ASDAS

Has been changed to:

3.2.22.4 Referenced based multiple imputation of ASDAS

Change #10

3.2.22.4 Referenced based multiple imputation of ASDAS

The following seeds have been added:

(seed will be 221).

(seeds will be 51, 52, 53, 54, 55, 56, 57, 58, 59, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 221).

Change #11

3.2.23 Time to flare

The following parts have been added:

- Subjects who discontinue the study without Withdrawal Visit will be treated as having experienced a flare at the next planned ASDAS visit. The final available ASDAS score date + 32 (4 Weeks + visit window) will be used as flare date. If Week 48 or Week 50 are last ASDAS visits, Week 48/Week 50 ASDAS date + 18. (2 weeks + visit window) will be used as flare date.
- …, if the last ASDAS visit date is not equal with the withdrawal date. If the withdrawal date is equal with the with the last ASDAS visit date, the same approach as for subject without Withdrawal Visit will be used for deriving the flare date.

…

  - If the date of the second missing ASDAS visit is available, the date of the visit will be used.
  - If the date of the second missing ASDAS visit is not available, the last ASDAS date + 60 (8 weeks +visit window) will be used as flare date. If Week 48 is the last ASDAS visits, Week 48 ASDAS date + 32. (4 weeks + visit window) will be used as flare date. If Week

50 is the last ASDAS visits, Week 50 ASDAS date + 46. (6 weeks + visit window) will be used as flare date.

Change #12

3.2.23 Time to flare

The following overview has been added:

Overview of different approaches for flare definition:

Table 12.2 Flare Definition and Time to Flare

Reason for flare	**Event/Censoring day**	**Non-responder imputation**	**Observed Case**	**Alternative handling of premature withdrawal**
ASDAS >=2.1 at 2 consecutive visits	Day of 2nd ASDAS	Event	Event	Event
ASDAS >3.5	Day of ASDAS	Event	Event	Event
Subject with Missing ASDAS at two consecutive visits	Planned date for the 2nd ASDAS visit	Event	Censored	Event
Subject discontinued study during Part B due to AE or Lack of efficacy	Day of discontinuation not equal ASDAS visit	Event	Censored	Event
Subject discontinued study during Part B due to other reason	Day of discontinuation not equal ASDAS visit	Event	Censored	Censored
No Flare (completed Part B)	Week 96	Censored	Censored	Censored
Incorrectly 'flared' due to delay in data entry	2nd ASDAS visit or planned date for the 2nd ASDAS visit if missing	Event	Censored	Censored

Change #13

3.2.26 Anti-CZP antibody status

Determination of ADAb will be done using a validated screening, confirmation, and titration ADAb bridging assay. The immunogenicity data will be analyzed according to dedicated Bioanalytical Analysis plans.

Has been changed to:

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

Change #14

3.2.27 Modified New York criteria

Subjects will be defined as having AS or non-radiographic axial spondyloarthritis (nr-axSpA) in this study based solely on the radiologic criterion. If a subject has sacroiliitis grade ≥2 bilaterally or sacroiliitis grade 3 to 4 unilaterally on x-ray then the subject will be considered to have AS/be modified New York (mNY) positive, otherwise they will be considered to have nr-axSPA/be mNY negative.

The mNY classification will determined at Screening. MRI will be assessed centrally and by 2 independent readers. If the results from the readers are different, then the result of the adjudicator will be used.

Has been changed to:

3.2.27 Modified New York criteria

Subjects will be defined as having radiographic axial spondyloarthritis (r-axSpA) or non-radiographic axial spondyloarthritis (nr-axSpA) in this study based solely on the radiologic criterion. For simplification, r-axSpA will be referred to AS in the SAP and statistical outputs. If a subject has sacroiliitis grade ≥2 bilaterally or sacroiliitis grade 3 to 4 unilaterally on x-ray then the subject will be considered to have AS/be modified New York (mNY) positive, otherwise they will be considered to have nr-axSPA/be mNY negative.

The mNY classification will be determined at Screening. MRI will be assessed centrally and by 2 independent readers. If the results from the readers are different, then the result of the adjudicator will be used as the adjudicator's decision was considered final for the inclusion of the subject in the study. If the modified NY criteria data recorded in the 'Modified NY Criteria for Ankylosing Spondylitis' eCRF differs from the MRI result reported by central review, the result from central review will be used in summary table instead of the result reported in the eCRF.

Change #15

The following section has been added:

3.2.28 Modified ASAS Classification Criteria

If the results for CRP, human leukocyte antigen B27 (HLA-B27) or MRI on the 'Modified ASAS Classification Criteria for AS' eCRF differ from central laboratory results and/or the MRI result from central review, respectively, the result of the central laboratory and/or central review of MRI will be used in the summary table for modified ASAS classification criteria

Change #16

3.3 Analysis time points

The following parts have been added:

If a subject withdraws early during Part A but does not have a Withdrawal Visit in Part A, safety and efficacy data recorded at either the date of the last scheduled or unscheduled visit during the treatment Period, whichever is later, will be included in summaries of "Early Withdrawal" visit. Only post Part A Baseline assessments will be included in the "Early Withdrawal" visit.

…

For subjects experiencing a flare, the last visit prior to escape medication (e.g. Escape Week 0) will be included in the summaries of the "Last/Withdrawal Visit". Only post Part B Baseline assessments will be included in the "Last/Withdrawal Visit".

…

Only post Flare baseline assessments will be included in the "Last/Withdrawal Escape" visit.

…

PK and ADA

For ADA data the same approach as for the efficacy analysis will be used, i.e for withdrawals to assign the withdrawal value to the next scheduled visit for the assessment. However, for PK analysis the withdrawal PK result will be only included in summaries of "Withdrawal" and the "Withdrawal Escape" and not assigned to the next scheduled visit. Furthermore, the Withdrawal summaries visit will not summarize any last visit assessments.

In addition, in the case that a sample is collected one or more days following the scheduled visit date, in which the drug was administered, the PK results for that sample will not be summarized as part of that visit.

Change #17

3.4 Definition of Baseline values

The following parts have been added:

If a Part A Baseline measurement is missing or not collected, and a Screening value is available, the Screening value will be utilized as Part A Baseline instead. If no measurement is available prior to receiving study medication in Part A, the Part A Baseline value is treated as missing.

…

If a Part B Baseline measurement is missing or not collected, and a post-Baseline Part A value is available, the post-Baseline Part A value will be utilized as Part B Baseline instead. If no post-Baseline Part A measurement is available prior to randomization date, the Part B Baseline value is treated as missing. If a subject does not receive study medication in Part B, the last available post-Baseline Part A measurement will be used as Part B Baseline.

…

If a Flare Baseline measurement is missing or not collected, and a post-Baseline Part B value is available, the post-Baseline Part B value will be utilized as Flare Baseline instead. If no post-Baseline Part B measurement is available prior to receiving study medication in Escape Part, the Flare Baseline value is treated as missing. If a subject does not receive study medication in Escape Part, the last available post-Baseline Part B measurement will be used as Flare Baseline.

…

When the date of first dose is derived, it should be based on the first injection of study treatment, regardless of whether it is an active treatment.

Change #18

3.4 Definition of Baseline values

For subjects randomized into Part B, a "Part B Baseline" will be defined as the latest measurement for that subject up to and including the day of administration of first study medication in Part B, unless otherwise stated

Has been changed to:

For subjects randomized into Part B, a "Part B Baseline" will be defined as the latest measurement for that subject up to and including the day of randomization, unless otherwise stated.

Change #19

The following Section has been removed:

3.6.8 Pharmacokinetic Set A

The Pharmacokinetic Set A (PKSA) will consist of all subjects from the SS who provide at least 1 PK sample during Part A.

Change #20

3.6.9 Flared Set

The Flared Set (FS) will consist of all subjects from the RS experiencing a flare in Part B. Since IXRS will be used to determine whether flare criteria are met during the course of the study, the population will be based on the IXRS data.

Has been changed to:

The Flared Set (FS) will consist of all subjects from the RS experiencing a flare (ASDAS>3.5 once or >2.1 on 2 consecutive visits, or ASDAS missing on 2 consecutive visits) in Part B. Since IXRS will be used to determine whether flare criteria are met during the course of the study, the population will be based on the IXRS data and whether they enter the Escape Part according to IXRS. **Change #21**

3.8 Center pooling strategy

The following country has been removed from Table 3.4: Canada.

Change #22

3.10 Changes to protocol-defined analyses

The protocol defines the censor date for time to flare as week 96. The censor date in the SAP is defined as the last possible visit of event (i.e. Week 92, or Week 94 if unscheduled occurred).

In addition to the protocol, a modified version for ASDAS-MI is added for sensitivity analysis of ASDAS-MI.

Has been changed to:

The SAP specifies which Baseline value is applicable for each analysis. The Flare Baseline is defined as the latest measurement for that subject up to and including the day of administration of first escape treatment in Part B instead of the visit in which the flare occurred.

In addition to the protocol, a modified version for ASDAS-MI is added for sensitivity analysis of ASDAS-MI.

Furthermore, analysis for IBD-Q response and IBD-Q remission using Part A Baseline and Part B Baseline as a reference and SF-36 PCS response and SF-36 MCS response using Part A Baseline as a reference are added.

Exploratory statistical comparisons for WPS will be based on MMRM instead of nonparametric bootstrap-t method.

ASDAS-Moderate Disease changed to ASDAS-Low Disease activity (Machado et al. 2018).

LOCF analysis are removed from Part B efficacy tables.

MMRM will be used instead of ANCOVA based on LOCF for

- Total spinal pain
- Nocturnal spinal pain
- PtGADA
- Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
- SF-36 PCS, MCS, and physical function domain
- Fatigue NRS
- ASQoL.

The primary purpose of the study is to evaluate treatment options of axSpA patients after being in sustained remission. Hence, one of the objectives is to evaluate the PK and immunogenicity of these patients. The CZP plasma concentration and ADAb titer of the patients that did not reach sustained remission are therefore not analysed, the Pharmacokinetic Set A (PKSA) as described in the protocol has been removed from the SAP.

Change #23

4.2 Handling of dropouts or missing data.

The following part has been removed:

If the Part A Baseline value is missing, carry forward the Screening value as far as Part A Baseline but no further.

Change #24

4.2 Handling of dropouts or missing data.

The following parts have been added:

The following continuous endpoints in Part B will be compared between treatment groups using a mixed model for repeated measures (MMRM). The pattern of missingness for these variables is assumed to be MAR. Subjects who enter Part B will have already demonstrated sustained remission through 48 weeks on CZP 200mg Q2W. Many of these subjects will be randomized to receive placebo and will, therefore, be more likely to experience a flare. Postflare data for subjects who experience a flare (in any treatment group) will be treated as missing for the MMRM analysis. It is further assumed that most subjects who discontinue at this late stage in the study (ie, during Part B) will do so as a result of reduced efficacy. Therefore, missing efficacy

data due to either flare or study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. ASDAS

- BASDAI
- Average of Questions 5 and 6 of the BASDAI concerning morning stiffness
- Fatigue item of BASDAI
- BASFI
- BASMI
- Total spinal pain score
- Nocturnal spinal pain score
- ASQoL
- PtGADA
- SF-36 physical functioning score
- SF-36 PCS
- SF-36 MCS
- WPS

…

The below rules will be applied for imputation of partial and missing medical history start dates:

- If start day is missing, but month and year are present, start date will be replaced by the first day of the month of that year or by the day of start of study medication in Part A, if study medication in Part A was first administered in that month and year (exception: if in this case start date is after stop date, also the first day of the month will be utilized).
- If start day and month is missing for one item diagnosis (medical history of any of the preferred terms ‘Axial Spondyloarthritis’, 'Spondyloarthropathy', 'Ankylosing spondylitis' , or 'Spondylitis') or symptom (medical history preferred term of ‘Back pain’, ‘Inflammatory pain’, ‘Spinal Pain’, ‘Lumbar spine pain’) and the year is similar, the missing month will be imputed by the month of the other item.
- If only one item diagnosis or symptom is present and if start day and month are missing and year is present, start date will be replaced by the first of January of that year or by the day and month of start of study medication in Part A, if study medication in Part A was first administered in that year (exception: if in this case start date is after stop date, also the first day of January will be utilized).
- If start day, month, and year are missing, start date will be replaced by the date of first study drug administration in Part A or by the first day of January of the year of the end date, if the first administration date is after the end date.

Change #25

4.3 Interim analyses and data monitoring

The following part has been added:

The following variable will be classed as unblinding during the full study. Therefore, these variables will be kept blinded until study unblinding after database lock:

- Randomized treatment group
- CRP (Post Week 48)
- MRI Efficacy scores (Post Week 48)
- PK and ADA (Post Week 48).

Change #26

4.7 Examination of subgroups

- Baseline CRP level (≤15 mg/L, >15 mg/L)
- ADAb status

Has been changed to:

- Baseline CRP level (≤ULN, >ULN)
- ADAb status (Negative, Positive <512, Positive 512-<1024, Positive >=1024).

Change #27

5.1 Subject disposition

The number and percentage of subjects who entered the study, completed Part A, discontinued from Part A with reasons for discontinuation, were randomized into Part B, completed Part B (without flare, after flare) and discontinued Part B with reasons for discontinuation will be presented.

…

- A listing of reasons for exclusion from the PPS

Has been changed to:

The number and percentage of subjects who entered the study, completed Part A, discontinued from Part A with reasons for discontinuation, were randomized into Part B, completed Part B (without flare, after flare) and discontinued Part B as well as started Escape Part, completed Escape Part and discontinued Escape Part with reasons for discontinuation will be presented.

…

- A listing of important protocol deviations (reasons for exclusion from the PPS)

Change #28

5.1 Subject disposition

The following part has been added:

Subjects who completed Week 48 without SFU are considered as completed Part A. Subjects completed Week 96 in Part B or in Escape Part without SFU visit are considered as completed Part B or completed Escape Part, respectively.

Subjects are considered to have started Escape Part, if they experienced a flare and have been moved to escape treatment via IXRS system (even if they do not receive any escape treatment).

…

The following listings will be produced based on the OLS

Change #29

6.1 Demographics

The following part has been added:

and region and country as per Table 3–4

Change #30

6.3 Medical history and concomitant diseases

Time since diagnosis of disease will be defined as: Earliest start date of the medical history of any of the preferred terms ‘Axial Spondyloarthritis’, 'Spondyloarthropathy' , 'Ankylosing spondylitis' , or 'Spondylitis' minus date of first study medication administration.

…

Date of first study medication administration – Start date of symptoms

Has been changed to:

Time since diagnosis of disease will be defined as: Date of first study medication administration minus earliest start date of the medical history of any of the preferred terms ‘Axial Spondyloarthritis’, 'Spondyloarthropathy', 'Ankylosing spondylitis', or 'Spondylitis'.

…

(Date of first study medication administration – Start date of symptoms)/365.25

Change #31

7 MEASUREMENTS OF TREATMENT COMPLIANCE

The CR ranges between 0 and 1. Specific details are provided in Section 12.3.

Compliance based on these methods will be summarized by axSpA subpopulation and overall in Part A using the SS, by treatment group in Part B using the SSB and by treatment group in Part B Post-flare using the ETS.

Has been changed to:

The CR ranges between 0 and 1. Specific details are provided in Section 12.2.

Compliance based on these methods will be summarized by axSpA subpopulation and overall in Part A using the SS, by treatment group in Part B using the SSB and by treatment group in Part B Post-flare using the ETS. The compliance for Part B will be calculated based on the first

injection in Part B and the compliance for Part B post-flare will be calculated based on the first injection in Escape Part.

If the Week 48 injection was not performed, the date of the Week 48 visit will be used as reference point for calculating the compliance of the first injection in Part B. Injections performed after the first injection in Part B will use the date of first injection in Part B as reference point for calculating the compliance of these injections.

If no Escape Week 0 injection was performed, the date of the Escape Week 0 visit will be used as reference point for calculating the compliance of the first injection in Escape Part. Injections performed after the first injection in Escape Part will use the date of first injection in Escape Part as reference point for calculating the compliance of these injections.

A by-subject listing of compliance will be provided by treatment group and axSpA subpopulation based on the SS. List of abbreviation

Change #32

8 EFFICACY ANALYSES

The following part has been added:

Furthermore, for selected analyses, statistical appendices will be created and included in the CSR appendix “Documentation of Statistical Methods”.

Change #323

8.1.1 Primary analysis of the primary efficacy variable

The number and percentage of subjects in Part B who do not experience a flare will be summarized descriptively for the RS using NRI data. Summary statistics will be presented by axSpA subpopulation and overall.

Has been changed to:

The 95% CI based on the Exact Binomial will also be presented. Summary statistics will be presented by treatment group in Part B.

Change #34

8.1.1 Primary analysis of the primary efficacy variable

The following part has been added:

To avoid the problem of the monotone likelihood resulting in infinite large confidence intervals (eg, if one of the cell counts in the 2x2 table is equal to zero), a penalized maximum likelihood approach based on the modified score procedure of Firth (eg, Heinze and Schemper, 2002) will be used in the logistic models.

Change #35

8.1.2 Secondary analyses of the primary efficacy variable

The following part has been added:

Factors for subgroups and the interaction term of subgroup and treatment will be added to the primary model specified in Section 8.1.1. The 95% CI based on the Exact Binomial distribution will also be presented.

The following part has been removed:

The region-by-treatment interaction will be tested by adding it to the primary model specified in Section 8.1.1. If the region-by-treatment interaction is found to be significant (α=0.10), further analyses will be conducted to determine which region or regions may be the source of the interaction. This will be done by running the logistic regression model (including the interaction term) where each region will be systematically removed from the model. The impact of a given region will be based on the change in the interaction p-value when that region is removed. The region or regions that appear to be driving the significant interaction effect will then be removed from the primary efficacy model to verify that conclusions remain the same with or without the influential region(s). Separately, a similar testing approach will be conducted to assess the mNY classification-by-treatment interaction. These analyses will not be part of the TFLs included in the CSR; however it will be included in the CSR appendix "Documentation of Statistical Method".

Change #36

8.1.3 Supportive and sensitivity analyses of the primary efficacy variable

The number and percentage of subjects in Part B who do experience a flare will be summarized descriptively for the RS using NRI data, indicating the contributing event (ASDAS defined flare, missing ASDAS values, premature study discontinuation).

The analysis described in Section 8.1.1 will be repeated for the RS using the different approaches as defined in Section 3.2.22 (observed case, handling of premature withdrawal, MMRM imputation, multiple imputation).

In addition, the analysis as described in Section 8.1.1 will also be repeated for the FAS and PPS using NRI data.

Has been changed to:

The number and percentage of subjects in Part B who do experience a flare will be summarized descriptively for the RS using NRI data, indicating the contributing event (ASDAS defined flare using scheduled visits only or using unscheduled visits, missing ASDAS values, premature study discontinuation with reason "AE or Lack of efficacy" or other reason, incorrectly 'flared' due to delay in data entry).

The flare by IXRS, calculated flare, time to flare and reason for flare according to NRI analysis will be listed based on the RS.

The analysis described in Section 8.1.1 will be repeated for the RS using the different approaches as defined in Section 3.2.22 (observed case, premature withdrawal, MMRM imputation, referenced-based multiple imputation). The 95% CI based on the Exact Binomial distribution will also be presented.

In addition, the analysis as described in Section 8.1.1 will also be repeated for the FAS and PPS using NRI data.

Change #37

8.2.1.1 Percentage of subjects achieving sustained remission at Week 48

The following part has been added:

The 95% CI based on the Exact Binomial distribution will also be presented.

Change #38

8.2.2 Part B

All tables will be presented for all subjects combined as well as broken out by axSpA subpopulation.

Has been changed to

All tables will be presented by treatment group in Part B.

Change #39

8.2.2.1 Time to flare

Kaplan-Meier plots of the time to flare will also be produced. Kaplan-Meier estimates of percentages of subjects who flare by Week 60, Week 72, Week 84 and Week 92 will be provided along with the corresponding 95% CI by treatment arm. For the calculation of the 95% CI, the LOGLOG transformation will be used.

These analyses will be repeated using the different approaches as defined in Section 3.2.23 (observed case, handling of premature withdrawal).

Has been changed to:

Kaplan-Meier plots of the time to flare will also be produced. Kaplan-Meier estimates of percentages of subjects who experience a experiencing a flare by Week 52, Week 56, Week 60, Week 64, Week 68, Week 72, Week 76, Week 80, Week 84, Week 88, Week 92 and Week 96 will be provided along with the corresponding 95% CI by treatment arm. For the calculation of the 95% CI, the LOGLOG transformation will be used.

These analyses will be repeated using the different approaches as defined in Section 3.2.23 (observed case, premature withdrawal).

Change #40

8.2.2.2 Continuous secondary efficacy variables

Summary statistics of observed and change from Part A Baseline to Week 96 values for ASDAS, BASDAI, BASFI, and BASMI will be presented by treatment group. These summaries will be based on observed case data and will be repeated using LOCF data.

The changes from Part A Baseline to Week 96 in ASDAS, BASDAI, BASFI and BASMI will be compared between treatment groups using a MMRM. The pattern of missingness for these variables is assumed to be missing at random (MAR). Post-flare data for subjects who experience a flare (in any treatment group) will be treated as missing in the MMRM analysis. The MMRM methods will be applied to all observed change from Part A Baseline (Week 0) data obtained from Week 48 up to Week 96.

…

Summary statistics of observed and change from Part A Baseline to Week 96 values for SIJ SPARCC and Spine ASspiMRI-a in the Berlin modification scores will be presented by treatment group.

The changes from Part A Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores at Week 96 will be compared between treatment groups using an analysis of covariance (ANCOVA) model. The model will include Part B Baseline score, treatment group, region, and mNY classification. Summary tables will present the adjusted means for PBO and the CZP doses, the respective differences to PBO, the corresponding p-values, and 95% CIs. This analysis will be performed using observed case data only.

Has been changed to:

Summary statistics of observed and change from Part B Baseline to Week 96 values for ASDAS, BASDAI, BASFI, and BASMI will be presented by treatment group. These summary statistics will be repeated for Part A Baseline.

The changes from Part B Baseline to Week 96 in ASDAS, BASDAI, BASFI and BASMI will be compared between treatment groups using a MMRM. The pattern of missingness for these variables is assumed to be missing at random (MAR). Post-flare data for subjects experiencing a flare (in any treatment group) will be treated as missing in the MMRM analysis. The MMRM methods will be applied to all observed change from Part B Baseline data obtained from Week 48 up to Week 96.

…

The comparison between treatment groups using MMRM will be repeated for Part A Baseline with Part A Baseline score as a fixed-effect covariate, treatment group, region, mNY classification, and visit as fixed-effect categorical factors, and Part A Baseline-by-visit and treatment group-by-visit as interaction terms.

Summary statistics of observed and change from Part B Baseline to Week 96 values for SIJ SPARCC and Spine ASspiMRI-a in the Berlin modification scores will be presented by treatment group. These summaries will be based on observed case data.

The changes from Part B Baseline in SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores at Week 96 will be compared between treatment groups using an analysis of covariance (ANCOVA) model. The model will include Part B Baseline score, treatment group, region, and mNY classification. Summary tables will present the adjusted means for PBO and the CZP doses, the respective differences to PBO, the corresponding p-values, and 95% CIs. This analysis will be performed using observed case data only.

These summary statistics will be repeated for Part A Baseline and the model includes Part A Baseline score, treatment group, region, and mNY classification.

Change #41

8.2.2.3 Binary secondary efficacy variables

These summaries will be based on observed case data and will be repeated using LOCF data (for ASDAS disease activity levels only) and using NRI data (for all other variables); improvement will be relative to Part A baseline only.

Has been changed to:

These summaries will be based on observed case data and will be repeated using NRI data (for all variables except ASDAS disease activity level); improvement will be relative to Part A Baseline only

Change #42

8.2.3 Subjects experiencing a flare in Part B

8.3.3 Subjects experiencing a flare in Part B

The following part has been added:

If deemed needed, further sensitivity analyses may be conducted excluding subjects incorrectly 'flared'.

Change #43

8.2.3.1 Continuous secondary efficacy variables

The following part has been added:

These summary statistics will be repeated for Flare Baseline. List of abbreviation.

Change #44

8.2.3.2 Binary secondary efficacy variables

The following part has been removed:

For comparative purposes, the tables will also include the responder number and percentages at the visit where flare criteria were met.

Change #45

8.3.1.4 SIJ SPARCC and spine ASspiMRI-a in the Berlin modification scores

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the SIJ SPARCC and Spine ASspiMRI-a in the Berlin modification scores for each visit by axSpA subpopulation and for all subjects combined. The analysis will use the average of the scores from the 2 independent reviewers. Additionally, agreement between readers will be assessed by calculation of intraclass correlation coefficients. These summaries will be produced for the OLS using observed case data only.

Has been changed to:

Summary statistics of the actual values and change from Part A Baseline values will be used to summarize the SIJ SPARCC and Spine ASspiMRI-a in the Berlin modification scores for each visit by axSpA subpopulation and for all subjects combined. The analysis will use results of the single read assessments for Part A from the independent reviewer. These summaries will be produced for the OLS using observed case data only.

Change #46

8.3.1.5 Individual ASAS core components

These summaries will be produced for the OLS using observed case data and will be repeated using LOCF imputation for missing data.

Has been changed to:

These summaries will be produced for the OLS using observed case data and will be repeated for BASFI and PtGADA using LOCF imputation for missing data.

Change #47

The following part has been updated for all the below sub sections.

8.3.1.6 PhGADA

8.3.1.11 Ankylosing Spondylitis Quality of Life (ASQoL)

8.3.1.12 Work Productivity Survey (WPS)

8.3.1.13 EuroQoL Health Status Questionnaire 5 dimensions 3 levels (EQ-5D-3L)

8.3.1.14 Maastricht Ankylosis Spondylitis Enthesitis Score (MASES)

8.3.1.15 Swollen and tender joint counts (44 joint count)

8.3.1.16 Spinal mobility

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

These summaries will be produced for the OLS using observed case data and will be repeated using LOCF imputation for missing data.

Has been changed to:

These summaries will be produced for the OLS using observed case data.

Change #48

8.3.1.15 Swollen and tender joint counts (44 joint count)

The following part has been removed:

In addition, the summaries for TJC will be repeated including only the subjects with a TJC>0 at Part A Baseline. Similarly, the summaries for SJC will be repeated including only the subjects with a SJC>0 at Part A Baseline.

Change #49

8.3.2.3 ASDAS, BASDAI, BASFI and BASMI

Summary statistics of the actual values, change from Part A Baseline values and change from Part B Baseline values for ASDAS, BASDAI, BASFI and BASMI will be presented by visit and treatment group. These summaries will be produced using observed case data and will be repeated using LOCF imputation for missing data.

The changes from Part A Baseline to each visit in Part B for ASDAS, BASDAI, BASFI and BASMI will be compared between treatment groups using a MMRM.

This analysis will be repeated based on change from Part B Baseline.

Has been changed to:

Summary statistics of the actual values, change from Part B Baseline values and change from Part B Baseline values for ASDAS, BASDAI, BASFI and BASMI will be presented by visit and treatment group. These summaries will be produced using observed case data. These summary statistics will be repeated for Part A Baseline.

The changes from Part B Baseline to each visit in Part B for ASDAS, BASDAI, BASFI and BASMI will be compared between treatment groups using a MMRM.

…

This analysis will be repeated based on change from Part A Baseline with Part A Baseline score as a fixed-effect covariate, treatment group, region, mNY classification, and visit as fixed-effect categorical factors, and Part A Baseline-by-visit and treatment group-by-visit as interaction terms.

Change #50

8.3.2.4 Individual ASAS core components

These summaries will be produced using observed case data and will be repeated using LOCF imputation for missing data.

For each component, the changes from Part B Baseline to each visit in Part B will be compared between treatment groups using a MMRM. (see Section 8.3.2.3). treatment group comparisons for both CZP groups vs. PBO will be performed based on the change from Part A Baseline values at the time points of Part B using an ANCOVA model including terms for Part B Baseline value, treatment group, region, and mNY classification. Tables will present the adjusted means for PBO and the CZP doses, the respective differences to PBO, the corresponding p-values, and 95% CIs. This analysis will be performed using LOCF imputation for missing data. This analysis will be repeated based on change from Part B A Baseline.

Has been changed to:

These summaries will be produced using observed case data.

For each component, the changes from Part B Baseline to each visit in Part B will be compared between treatment groups using a MMRM. (see Section 8.3.2.3). This analysis will be repeated based on change from Part A Baseline.

Change #51

The following part has been updated for all the below sub sections.

8.3.2.5	Fatigue NRS
8.3.2.7	ASQoL
8.3.2.8	Work Productivity Survey (WPS)
8.3.2.11	Short-Form 36-Item Health Survey (SF-36)

This summary will be produced using observed case data and will be repeated using LOCF imputation for missing data.

Treatment group comparisons for both CZP groups vs. PBO will be performed based on the change from Part A Baseline in Fatigue using an ANCOVA model including terms for Part B Baseline score, treatment group, region, and mNY classification. Tables will present the adjusted means for PBO and the CZP doses, the respective differences to PBO, the corresponding p-values, and 95% CIs. This analysis will be performed using LOCF imputation for missing data. This analysis will be repeated based on change from Part B Baseline.

Has been changed to:

This summary will be produced using observed case.

Treatment group comparisons for both CZP groups vs. PBO will be performed based on the change from Part B Baseline to each visit in Fatigue using a MMRM. (see Section 8.3.2.3). This analysis will be repeated based on change from Part A Baseline.

Change #52

8.3.2.5 Fatigue NRS

The following part has been added:

The remaining components of the BASDAI maybe analyzed as well.

Change #53

8.3.2.8 Work Productivity Survey (WPS)

These summaries will be produced using observed case data and will be repeated using LOCF imputation for missing data.

Change #54

8.3.2.9 Number of uveitis flares, IBD exacerbations and psoriasis exacerbations

These summary tables will display the number and percentage of subjects experiencing the considered event in Part B and number of event. Only the events occurring prior to ASDAS-defined flare, if ASDAS-defined flare occurred, will be considered. In this style of output, "number of event" will include all occurrences of the event including repeat occurrences in individual subjects, while "number of subjects" will count each subject only once.

Has been changed to:

These summary tables will display the number and percentage of subjects experiencing the considered event in Part B, and number of event, incidence rate, event rate and 100 patient exposure year. Only the events occurring prior to up to Escape Week 0 ASDAS-defined flare, if ASDAS-defined flare occurred, will be considered. In this style of output, "number of event" will include all occurrences of the event including repeat occurrences in individual subjects, while "number of subjects" will count each subject only once.

Change #55

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

This analysis will be performed using Part A Baseline as reference for the definition of SF-36 PCS and MCS response and will be repeated using Part B Baseline as reference.

Has been changed to:

This analysis will be performed using Part A Baseline as reference for the definition of SF-36 PCS and MCS response.

Change #56

8.3.2.12 Descriptive summaries of other efficacy variables

The below listed efficacy variables will be summarized descriptively by visit and by treatment group as applicable using observed case and LOCF imputed data based on the RS unless otherwise specified

Has been changed to:

The below listed efficacy variables will be summarized descriptively by visit and by treatment group as applicable using observed case based on the RS unless otherwise specified

Change #57

9.1 Pharmacokinetics

CZP plasma concentrations will be tabulated and summarized using PKSA (overall and by overall anti-CZP Ab status) for each visit at which samples were taken during Part A using the geometric mean, 95% CI, geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. Plasma concentration time curves will be plotted overall, and by overall anti-CZP Ab status.

CZP plasma concentrations will be tabulated and summarized by treatment group using PKSB (overall and by overall anti-CZP Ab status.) for each visit at which samples were taken during Part B using the geometric mean, 95% CI, geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. For each treatment group, plasma concentration time curves will be plotted overall, and by overall anti-CZP Ab status.

Part B summaries will not include post-flare assessments. Separate summaries will be produced for subjects who flare in Part B and start escape therapy including only the post-flare assessments. Concentrations below the limit of quantification (BLQ) will be replaced by the LLOQ divided by 2 for the calculation of descriptive statistics.

Subgroup analyses based on age, gender, race, region, CRP level and mNY classification will be performed.

Has been changed to:

CZP plasma concentrations will be tabulated and summarized using PKSB for each visit at which samples were taken during Part A by axSpA subpopulation using the geometric mean, 95% CI, geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. Plasma concentration time curves of Part A will be plotted for all subjects and Part A treatment-emergent ADAb (Part A TE ADAb) status with titer classification (see Section 9.2).

CZP plasma concentrations will be tabulated and summarized by treatment group using PKSB for each visit of the total study at which samples were taken using the geometric mean, 95% CI,

geometric CV, arithmetic mean, arithmetic SD, median, minimum, and maximum. The table will be repeated for TE ADAb status with titer classification (see Section 9.2).

For each treatment group, plasma concentration time curves of Part B will be plotted and repeated for TE ADAb status with titer classification (see Section 9.2).

Part B figures will not include post-flare assessments. Separate figures will be produced for subjects who experience a experiencing a flare in Part B and start escape therapy including only the post-flare assessments.

Spaghetti plots of CZP plasma concentrations by week from CZP first dosing separated by treatment group and TE ADAb status and titer classification (see Section 9.2) will be presented for subjects with flare and no flare per NRI approach.

Spaghetti plots of CZP plasma concentrations by week from CZP first dosing separated by treatment group will be presented for flared subjects per NRI by TE ADAb status and titer classification (see Section 9.2).

Individual plots by subject of CZP Concentrations/ADAb titer and ASDAS score plotted on the Y-axes by visit (X-axis) for the full treatment period, including SFU. Flare data will be displayed on the days relative to Part A.

Concentrations below the limit of quantification (BLQ) will be replaced by the LLOQ divided by 2 for the calculation of descriptive statistics.

Anomalous values will be not included in summaries/analysis and will be reviewed and flagged by pharmacokineticist.

All CZP plasma concentration will be listed at subject-level.

Change #58

9.2 Immunogenicity

Immunogenicity will be assessed through listing of individual results by subject and summary tables. Immunogenicity data will be correlated with PK and efficacy readout. In addition, immunogenicity will be correlated with possible safety findings.

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 ADAb assay. Subjects will receive an overall classification, inclusive of Baseline and Post-Baseline results, and be classified as follows based on the ADAb assay results:.

- ADAb negative: no confirmed positive ADAb samples at any of the sampling time points.
- ADAb positive: confirmed positive ADAb samples at one or more sampling time points.

Summaries will be done separately in Part A by axSpA subpopulation and overall using PKSA and in Part B by treatment group using PKSB. Part B summary will not include post-flare assessments. Separate summary by treatment group will be produced for subjects who flare in Part B and start escape therapy including only the post-flare assessments.

Has been changed to:

Immunogenicity will be assessed through listing of individual results by subject and summary tables. Immunogenicity data will be correlated with PK and efficacy readout.

A cut point will be determined by the bioanalytical laboratory during assay validation. This cut point will be used to determine the status of ADAb in the test sample.

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 positive screen and positive immunodepletion.
- An ADAb status of negative will be concluded for any sample with an ADAb level that is either negative screen or positive screen and negative immunodpepletion.

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 ADAb assay. ADAb titer dilution factor is presented in the listings and summaries excluding the minimum required dilution (MRD), i.e. reportable values are divided by 100.

If the titer for an ADAb level that is positive screen and positive immunodepletion is missing, then a conservative will be used and ADAb status will be consider as positive. No imputation rules apply for the missing titer.

If the ADAb level is positive screen but no confirmatory result could be determined, then then a conservative will be used and ADAb status will be consider as positive.

Anomalous value will be not included in summaries/analysis and will be reviewed and flagged by pharmacokineticist.

Subjects will have the following Baseline ADAb status:

Baseline ADAb Status positive is defined as having an ADAb confirmed positive inhibition result at baseline, Baseline ADAb Status negative is defined as having negative inhibition result at baseline.

Subjects will receive a treatment-emergent ADAb status, inclusive of Baseline and Post-Baseline results (including SFU), and be classified as follows based on the ADAb assay results:

- TE ADAb status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥minimum significant ratio (MSR) -fold

increase from baseline on CZP treatment. The MSR will be defined during the process of sample analysis and is disease-specific.

- TE ADAb status negative is defined as having no samples either ADAb positive or with values ≥MSR -fold increase from baseline.

Once determined positive, the highest titer during Part A and Part B (including Escape and SFU) is used to categorize the subject.

- Positive <512,
- Positive 512-<=1024,
- Positive >1024.

In addition, for some outputs also the subjects with TE ADAb status negative will be presented.

The TE ADAb will be also defined for the different treatment periods. For each Period the SFU visit will be considered:

- Part A TE ADAb status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline during Part A or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥MSR increase from baseline on CZP treatment during Part A. Once determined positive, the highest titer during Part A is used to categorize the subject.
- Part A TE ADAb status negative is defined as having no samples either ADAb positive or with values ≥MSR increase from baseline during Part A.
- Part B TE ADAb Status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline during Part B or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥MSR increase from baseline on CZP treatment during Part B. Once determined positive, only the highest titer during Part is used to categorize the subject. For subjects with flare in Part B, any assessments after Escape treatment are not included.
- Part B TE ADAb Status negative is defined as having no samples either ADAb positive or with values ≥MSR increase from baseline during Part B (excluding any values up to 8 weeks post first Part B dose).
- Escape TE ADAb Status positive is defined as either (i) baseline ADAb negative subjects having at least one ADAb confirmed positive sample post baseline during Escape Part or ii) baseline ADAb positive subjects with at least one post baseline sample with ≥MSR increase from baseline on CZP treatment during Escape Part B. Once determined positive, only the highest titer during Escape is used to categorize the subject.
- Escape TE ADAb Status negative is defined as having no samples either ADAb positive or with values ≥MSR increase from baseline during Escape Part. Once determined positive, only the highest titer during Escape is used to categorize the subject.

Summaries will be done using PKSB.

Summary of shift from Baseline ADAb status with titer classification to TE ADAb Status with titer classification by treatment group and for all subjects will be presented during the entire study.

The shift table will be repeated for

- Part A TE ADAb status with titer classification to Part B TE ADAb status with titer classification by treatment group and for all subjects during Part B.
- Part B TE ADAb status with Part B titer classification to Escape Part TE ADAb status with Escape titer classification by treatment group and for all subjects during Escape Part.

The time to achieving TE ADAb for all subjects will be analyzed based on Kaplan-Meier methods. Subjects will be considered to have an event at the time where treatment-emergent ADAb positive is first achieved during treatment period excluding Baseline/pre-treatment. Subjects classified as treatment-emergent ADAb negative will be censored at the time of last available ADAb result. The median and 95% CI based on the Kaplan-Meier estimation will also be presented. A plot of time to first ADAb positivity will be presented.

Number and percentage of subjects with ADAb titer above the specified cut point ("categorized") at any visit during the treatment period, separated by treatment group will be presented.

A summary of flare responder during Part B (NRI) and ASDAS MI responder at Week 96 (NRI) by treatment group and by highest ADAb titer ("categorized") per subject will be presented during Part A and Part B (excluding Escape). This table will be repeated for Part B (excluding Escape).

A scatter plot of CZP Plasma Concentration and ADAb titer during Part A for all subjects will be presented. The scatter plot will be repeated for Part B by treatment group.

Spaghetti plots of ADAb titer on a loglog scale by week from CZP first dosing separated by treatment group and TE ADAb status/titer classification will be presented for subjects with flare and no flare per NRI approach.

All individual ADAb results will be listed at subject level.

Change #59

10.1.2.1 All subjects in SSB – Prior to Flare

Exposure to study medication in Part B will be evaluated by summarizing the total number of PBO injections, the total number of CZP or PBO injections, the total number of CZP injections, the total amount of CZP received in mg, the total drug exposure duration in days, the patient time at risk and the total patient years at risk. Summaries will be based on the SSB and presented by treatment group.

Has been changed to:

Exposure to study medication in Part B will be evaluated by summarizing the total number of CZP or PBO injections, the total number of CZP injections, the total amount of CZP received in mg, the total drug exposure duration in days, the patient time at risk and the total patient years at risk. Summaries will be based on the SSB and presented by treatment group.

Change #60

10.2 Adverse events

Overall summaries of AEs will be produced to present the number of events and number and percentage of subjects with:

- Any TEAE
- Any serious TEAE
- Any TEAE leading to study discontinuation
- Any drug-related TEAE
- Any severe TEAE
- Deaths (TEAE leading to death)

In addition, the following AE summaries will be presented:

- All TEAEs by primary SOC, HLT, and PT
- All serious TEAEs by primary SOC, HLT, and PT
- All non-serious TEAEs by primary SOC, HLT, and PT
- All TEAEs leading to permanent discontinuation of study drug by primary SOC, HLT, and PT
- All TEAEs leading to death by primary SOC, HLT, and PT
- All TEAEs by primary SOC, PT, and maximum intensity
- All TEAEs by primary SOC, PT, and maximum relationship
- All serious TEAEs by primary SOC, PT and maximum relationship
- All non-serious TEAEs by primary SOC, PT and maximum relationship
- All TEAEs above reporting threshold of 5% by primary SOC and PT
- All non-serious TEAEs above reporting threshold of 5% by primary SOC and PT
- All non-serious TEAEs above reporting threshold of 5% by primary SOC, PT and maximum relationship
- All injection related TEAEs by primary SOC, HLT, and PT

AE summary tables will also be produced to display the exposure adjusted incidence rate (EAIR) with associated 95% CI, and the exposure adjusted event rate (EAER):

- All TEAEs by primary SOC, HLT, and PT
- All serious TEAEs by primary SOC, HLT, and PT

Has been changed to:

Overall summaries of AEs will be produced to present the number of events and number and percentage of subjects with:

- Any TEAE

- Any serious TEAE
- Any TEAE leading to study discontinuation
- Permanent withdrawal of study medication due to TEAEs
- Any drug-related TEAE
- Any severe TEAE
- All Deaths (AEs leading to death)
- Deaths (TEAE leading to death)

In addition, the following AE summaries will be presented:

- All TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All serious TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All TEAEs leading to permanent discontinuation of study drug by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All TEAEs leading to death by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c)
- All TEAEs by primary SOC, PT, and maximum intensity (for the Parts/Analysis sets a, b, c, d)
- All TEAEs by primary SOC, PT, and maximum relationship (for the Parts/Analysis sets a, b, c, d)
- All serious TEAEs by primary SOC, PT and maximum relationship (for the Parts/Analysis sets a, b, c, d)
- All TEAEs above reporting threshold of 5% by primary SOC and PT (for the Parts/Analysis sets a, b, c, d)
- All non-serious TEAEs above reporting threshold of 5% by primary SOC and PT (for the Parts/Analysis sets a, b, c)
- All injection related TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)

AE summary tables will also be produced to display the exposure adjusted incidence rate (EAIR) with associated 95% CI, and the exposure adjusted event rate (EAER):

- All TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)
- All serious TEAEs by primary SOC, HLT, and PT (for the Parts/Analysis sets a, b, c, d)

Change #61

10.2 Adverse events

For EAERs, the numerator will be the number of AEs including repeat occurrences in individual subjects; the denominator will be 100 patient-years. That is, the total summation of individual patient-years at risk divided by 100. No CI will be computed for EAER.

The following AE summary tables will be repeated by overall anti-CZP Ab status (positive/negative):

- Overall summary of AEs
- All TEAEs by primary SOC, HLT, and PT
- All serious TEAEs by primary SOC, HLT, and PT

The following AEs of interest will be summarized:

1. Serious infections, including opportunistic infections
2. Malignancies, including lymphoma
3. Serious cardiovascular events

Has been changed to:

The TEAE summary table will be presented by TE ADAb status and depending on the timepoint of onset of TEAE (Prior to becoming ADAb Positive>=1024, After becoming ADAb Positive>=1024, Subjects Who Never Became ADAb Positive>=1024) based on PKSB.)

The following AEs of interest will be summarized:

1. Serious infections, including opportunistic infections
2. Malignancies, including lymphoma
3. Serious cardiovascular events (i.e. Major adverse cardiac events or MACE)

Change #62

10.2 Adverse events

– Ischaemic central nervous system vascular conditions (SMQ)

Has been changed to:

– Ischaemic central nervous system vascular conditions (SMQ) except events coding to PT "Transient ischaemic attack"

Change #63

10.2 Adverse events

The following part has been added:

Injection reactions are identified by the investigator on the eCRF.

The following Category C consists of the following PTs has been added:

- Rash erythematous

Change #64

10.2 Adverse events

The following AE summary tables will be used as the basis for presentation of AEs of interest as well as for hepatic events, hypersensitivity reactions and anaphylactic reactions:

Has been changed to:

The following AE summary tables (for the Parts/Analysis sets a, b, c, d) will be used as the basis for presentation of AEs of interest as well as for hepatic events, hypersensitivity reactions and anaphylactic reactions:

Change #65

10.3 Clinical laboratory evaluations

Testing for hepatitis B surface antigen and antibodies to hepatitis C and HIV will be performed at Screening. Screening for HLA-B27 will be performed at Baseline only. These data will be listed only.

Has been changed to:

Testing for hepatitis B surface antigen and antibodies to hepatitis C and HIV will be performed at Screening. HLA-B27 will be performed at Baseline only. These data will be listed only.

Change #66

10.3 Clinical laboratory evaluations

Descriptive statistics for observed values and change from Baseline will be presented for each scheduled visit for the following parameters:

…

- Biochemistry: sodium, potassium, chloride, bicarbonate, total calcium, inorganic phosphorus, creatine phosphokinase (CPK), glucose, creatinine, uric acid, urea, total protein, albumin, alkaline phosphatase (AP), aspartate aminotransferase (AST), alanine aminotransferase (ALT), bilirubin, total cholesterol

Has been changed to:

Descriptive statistics for observed values and change from Part A Baseline will be presented for each scheduled visit for the following parameters:

…

- Biochemistry: sodium, potassium, chloride, bicarbonate, total calcium, inorganic phosphorus, creatine phosphokinase (CPK), glucose, creatinine, uric acidurate, urea nitrogen, total protein, albumin, alkaline phosphatase (AP), aspartate aminotransferase (AST), alanine aminotransferase (ALT), bilirubin, total cholesterol

Change #67

10.3 Clinical laboratory evaluations

The following parts have been added:

A by-subject listing of all laboratory data will be provided. This listing will be presented by treatment group, results (with abnormal values flagged) and unit. In the listing for biochemistry also Glomerular Filtration Rate will be added.

…

A by-subject listing of markedly abnormal values will be provided separately.

…

A by-subject listing for any subject with at least one post-baseline elevated test data will be provided. The listing includes all liver function test laboratory values for this subject.

Samples and results for additional laboratory tests (hepatitis B surface antigen and antibodies to hepatitis C and HIV, Genomics Laboratory Data) will be listed separately.

Change #68

10.4.1 Vital signs

Descriptive statistics for observed values and change from Baseline will be presented for each scheduled visit for each vital sign parameter.

Has been changed to:

Descriptive statistics for observed values and change from Part A Baseline will be presented for each scheduled visit for each vital sign parameter.

Change #69

10.4.2.1 Pregnancy testing

Pregnancy testing must be carried out for women of childbearing potential and will consist of serum testing at Screening and SFU, and urine testing at Baseline and Week 96/withdrawal Visit. Pregnancy test results will be listed only.

Has been changed to:

Pregnancy testing must be carried out for women of childbearing potential and will consist of serum testing at Screening and SFU, and urine testing at Baseline and Week 96/withdrawal Visit. Pregnancy test results (i.e. Choriogonadotropin Beta) will be listed only.

Change #70

10.4.2.2 Physical assessments

Summary statistics will be provided by axSpA subpopulation and overall in Part A using the SS and by treatment group in Part B using the SSB by visit presenting actual values and changes from Baseline for weight, height and BMI, as applicable.

Has been changed to:

Summary statistics will be provided by axSpA subpopulation and overall in Part A using the SS and by treatment group in Part B using the SSB by visit presenting actual values and changes from Part A Baseline for weight, height and BMI, as applicable.

Change #71

11 REFERENCES

The following reference has been added:

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.

The following reference has been removed:

Ware J.E., Jr., Kosinski M., Bjorner J.B., Turner-Bowker, D.M., Gandek, B. and Maruish, M.E. (2007) User's Manual for the SF-36v2TM Health Survey (2nd Ed). Lincoln, RI, QualityMetric Incorporated.

Change #72

The following Section has been removed and the following Sections receive a new number.

12.1 Nonparametric bootstrap-t method

Barber and Thompson note that it is a common statistical problem in the analysis of cost data from clinical trials to provide comparisons and inferences regarding the total costs associated with each treatment group, in the presence of highly skewed cost distributions (Barber and Thompson, 2000). Standard non-parametric tests (that compare the overall shape of distributions) or log transformation to normalize data (that compare geometric means) are not always suitable because for cost data the analysis needs to focus on the arithmetic mean costs rather than the overall differences in the shape of the distribution, the medians, or the geometric means. The relevance of the arithmetic mean is that it can be used to calculate the total costs of treatment, which is needed to inform health economic policy decision makers. The non-normality of cost distributions also means that standard parametric comparisons of means based on the normal distribution (ANOVA and two-sample t-tests) may not be suitable. In such cases, bootstrap techniques can be recommended when making inferences about arithmetic means for moderately sized samples of highly skewed data such as healthcare costs. The cost data approach will be applied to the WPS.

The Bootstrap principle

The bootstrap is a data based simulation method that is useful for deriving confidence intervals and hypothesis testing when the sampling distribution of an estimator is not known or cannot be defined mathematically.

A balanced bootstrap resampling method (Gleason, 1988) is used to increase the precision of the bootstrap bias and the standard error (SE), where each observation occurs a total of B times in the collection of B bootstrap samples. This does not force each bootstrap sample to contain all observations; the first observation may occur twice in the first bootstrap sample and not at all in the second, while the second observation may occur once in each sample.

Samples of the same size as the original sample used are drawn separately from each treatment group by sampling with replacement from the original data. For each resample the statistic of interest (mean or difference in means) is calculated. The distribution of these B (where B = number of bootstrap samples) values provides an approximation of its population sampling distribution and can be used to estimate confidence intervals and to conduct hypothesis tests. Typically, a large number of bootstrap resamples are calculated (10000 resamples in this analysis).

The bootstrap-t method involves generating bootstrap-t values, which can be used in place of the standard t-distribution values to calculate confidence intervals and p-values.

Bootstrap resampling is used to obtain an approximation to the distribution of the test statistic under the null hypothesis of equal means. This distribution is then directly compared with the observed value of the test statistic to estimate the p-value. The most reliable test statistic for comparison of means is the Studentized statistic, which is the difference in means divided by its standard error. This is calculated for each bootstrap resample as:

$$T_b^* = (M_b^* - M)/SE_b^*$$

where M_b^* is the difference in bootstrap means between treatment groups, M the observed difference in means between treatment groups, and SE_b^* the standard error of the bootstrap difference in means between treatment groups.

The standard error of the difference in means is calculated using the same method as for a t- test with unequal variances:

$$SE_b^* = \text{sqrt}\ [\ ((SD_{Yb}^*)^2/m) + ((SD_{Zb}^*)^2/n)\ ]$$

where SD_{Yb}^* and SD_{Zb}^* are the observed standard deviations of the bootstrap samples Y_b^* and Z_b^* in each treatment group (group Y with sample size m and group Z with sample size n).

Note that in some rare cases SEb*=0. When SEb* and $(M_b^* - M)$ are both equal to 0, then $T_b^* = (M_b^* - M)/SE_b^*$ is considered to be 0. When SEb*=0 and $(M_b^* - M) \neq 0$, then $T_b^* = (M_b^* - M)/SE_b^*$ is considered to be tending towards infinity and is, therefore, greater than tobs in the p-value calculation described in next Section, Bootstrap hypothesis test.

12.1.2 Bootstrap hypothesis test

The estimated distribution of the test statistic under the null hypothesis of no difference in means is used instead of a standard t-distribution to obtain the approximate two-sided p-value:

$$P\text{-value} \qquad \hat{P}_{boot} = \#\{|t_b^*| \geq |t_{obs}|\} / B,$$

where tobs is the observed value of the test statistics. If the distribution of bootstrap-t values is close to the standard t-distribution then the p-value from the usual t-test and that from the bootstrap-t will be similar.

It is recognized that the bootstrap-t test performs best on a variance-stabilized scale where the difference in means and its standard error are independent. This independence can be assessed from a plot of the bootstrap values (ie differences in means M_b^*) against their corresponding standard errors (SE_b^*). If there is a strong relationship between these, the test should be carried out after a variance stabilizing transformation.

The current analysis will be conducted after a variance stabilizing transformation of the bootstrap values. (See details on the variance stabilizing transformation in Section 12.1.4)

12.1.3 Bootstrap confidence intervals

To estimate a 100(1−α)per cent confidence interval, the 100(α/2)per cent and 100(1−α/2)per cent percentiles of T* ($T^*_{(\alpha/2)}$ and $T^*_{(1-\alpha/2)}$, respectively) are used.

For the estimation of a 100(1-α) per cent confidence interval for the difference in means, the confidence interval is given by:

$$(M - T^*_{(1-\alpha/2)}\,SE(M),\ \ M - T^*_{(\alpha/2)}\,SE(M))$$

where M is the observed difference in means and SE(M) its standard error.

The distribution of bootstrap-t values (T*) is obtained in the same way as previously described for hypothesis testing. The 100α percentile is estimated by the α(B+1)th member of the ordered bootstrap sample if this is a whole number; otherwise linear interpolation must be used (Davison and Hinkley, 1997).

As before for the bootstrap test, this method of obtaining a confidence interval is best performed on a variance stabilized scale. The confidence interval is thus calculated on the variance stabilized scale and back transformed to obtain an interval on the original scale. (See details on the variance stabilizing transformation in Section 12.1.4)

12.1.4 Variance stabilizing transformation

The bootstrap hypothesis test presented in Section 12.1.2 gives different results if a different scale is used for the bootstrap differences in means M_b^*, and will perform best on a variance stabilized scale where M and SE(M) are independent.

The test will be carried out after a variance stabilizing transformation of the bootstrap values M_b^*, say function ***g***. Note that this is not the same as transforming the data since transforming the bootstrap values M_b^* will still allow a comparison of arithmetic means.

The test statistic calculated for each resample is then given by

$$T_b^* = g(Mb^*) - g(M).$$

There is no longer a need to divide by the estimated standard error because on the variance stabilized scale this will be constant.

An estimate of the P-value for the test is then obtained by comparing the observed value of the test statistic

$$t_{obs} = g(M) - g(M0),$$

where M_0 is the null value for the difference in means, usually 0) with the distribution of T_b^* values.

If none of the values of M_b^* is equal to 0, then $g(M_0)$ is interpolated by

$$g(M_0) = \frac{g(M_{0-}) + g(M_{0+})}{2},$$

where M_{0+} is the smallest value of M_b^* that is greater than 0 and M_{0-} is the greatest value of M_b^* that is less than 0. In cases where M_{0+} does not exist, $g(M_0)$ is estimated by $g(M_{0-})$. In cases where M_{0-} does not exist, $g(M_0)$ is estimated by $g(M_{0+})$.

Details of an `automatic' method of finding an approximate variance stabilizing transformation ***g*** are given below. It was recommended to base this 'automatic' method on a Taylor series argument. For a random variable X with mean M and standard deviation s(M) that varies as a function of M, the transformation given by

$$g(x) = \int^{x} \frac{1}{s(u)}\,\mathrm{d}u \qquad (1)$$

has the property that the variance of g(X) is approximately constant.

The Non-Parametric Bootstrap-t Method with a variance stabilizing transformation

The following steps describe the variance stabilized non-parametric bootstrap-t method for a difference in means:

1. Form a set of B bootstrap data sets sampled with replacement from the two groups Y and Z, as described in Section **Error! Reference source not found.**. The number of r eplications B in this analysis is 10000.
2. Compute the difference in means M_b^* and standard error SE_b^* for each bootstrap data set, b=1, 2,..., B.
3. Fit a curve to the points [M_b^* and SE_b^*] using a non-linear regression technique to produce a smooth function s such that $s(M_b^*)$ is the average SE_b^* at M_b^*, using quadratic polynomials for the local regression and cubic interpolation polynomials in the blending method of the local polynomial fits. The generalized cross-validation smoothing criterion is applied for the fit.

   The size of the local neighborhood in the local fitting is determined by the value of the smoothing parameter (θ), which indicates the proportion of the data points used for the local fitting. The value of the smoothing parameter θ that is used, is obtained by minimizing the generalized cross-validation smoothing criterion. The cubic interpolation method assumes that the maximum number of points in the leaf nodes of the kd tree (bucket size) is set to θ*k/5, where θ is the estimated smoothing parameter and k is the number of observations being used.
4. Estimate the variance stabilizing transformation g(M) using equation (1) and a numerical integration technique. For the present analysis, the Simpson's numerical integration method will be used.
5. Compute a bootstrap-t confidence interval for the transformed values $g(M_b^*)$ in the way described previously in Part C, where α=0.05. The standard error will be approximately constant, so SE_b^* and SE(g(M)) can be set equal to 1. Thus the CI for the transformed values will be calculated as follows:

   $$(g(M) - T^*_{(1-\alpha/2)}, \quad g(M) - T^*_{(\alpha/2)}),$$

   where $T^*_{(1-\alpha/2)}$ and $T^*_{(\alpha/2)}$ are the percentiles of $T_b^* = g(M_b^*) - g(M)$.

   Estimate the p-value for the hypothesis test as described previously in Section **Error! R eference source not found.**.

6. The endpoints (ie, the estimated difference in means and the CI) calculated on the transformed scale can be mapped back to the original scale using the inverse transformation g^{-1}. Of note, when the value falls in between two existing values in the lookup table, the inverse transformation g^{-1} will be calculated using linear interpolation.

   (Programming note: For the CI, in the case where the CI exact limits cannot be determined because they fall outside the range of existing data, then the output will not display the limit which cannot be determined, but will display the following labels: "<min" or ">max", where applicable.)

   The adequacy of the transformation can be checked by plotting M_b^* against the standard error on the variance stabilized scale, an approximation of which is given using the delta method.

$$SE\ (g(M_b^*)) = SE_b^* / s(M_b^*)$$

Change #73

12.2 Compliance ratio calculation

- Exposure Duration [b] (days) = week 94 Visit/last injection date –Week 48 Visit date + 14

The cumulative difference will be calculated for all visits from Week 48 to week 94.

For the measurement of treatment compliance in Part B post-flare, the post-flare exposure duration, identified as "Exposure Duration [f]", will use the last of the injection escape Visit dates:

- Exposure Duration [f] (days) = last escape visit/last injection date –Escape Week 0 Visit date + 14

Has been changed to:

- Exposure Duration [b] (days) = week 94 Visit/last injection date – Week 48 injection/Week 48 Visit date + 14

The cumulative difference will be calculated for all visits from Week 48 to week 94.

For the measurement of treatment compliance in Part B post-flare, the post-flare exposure duration, identified as "Exposure Duration [f]", will use the last of the injection escape Visit dates:

- Exposure Duration [f] (days) = last escape visit/last injection date – first Escape Week 0 injection/Escape Week 0 Visit date + 14