Sandoz Biopharmaceuticals Clinical Development

GP2017 (INN: Adalimumab)

Clinical Trial Protocol GP17-302 / NCT02744755

A randomized, double-blind, parallel-group, multicenter study to demonstrate similar efficacy and to compare safety and immunogenicity of GP2017 and Humira® in patients with moderate to severe active rheumatoid arthritis

SAP – Detailed Statistical Methodology

Author(s):	[redacted]
Sponsor:	Hexal AG, Industriestr. 25, 83607 Holzkirchen, Germany
Date:	19-Mar-2018
Version:	Final 2.0
Number of pages:	43

[redacted]

Document History – Changes compared to previous version of the SAP

Version	Date	Changes
1.0	04-Aug-2017	Not applicable – First Version
2.0	19-Mar-2018	Per-protocol set definition was updated according to BDRM minutes. Definitions of PD classification (major/minor) were added according to BDRM minutes. Sensitivity analysis of primary endpoint and key secondary endpoint were added according to BDRM minutes. Appendix 3 (Adverse events of special interest) and Appendix 4 (Medications of special interest) were updated by medical team according to the AE and medication database.

[redacted]

Table of contents

Document History – Changes compared to previous version of the SAP 2
Table of contents 3
List of tables 4
List of figures 4
1 Introduction 7
2 Statistical and analytical plans 7
2.1 Study documents and general considerations 7
2.1.1 Study documents 7
2.1.2 General considerations 8
2.2 Study objectives and endpoints 9
2.3 Assessments 11
2.3.1 Efficacy assessments 11
2.3.2 Safety and tolerability assessments 17
2.4 Statistical method planned in the protocol 20
2.4.1 Analysis of the primary endpoint 20
2.4.2 Analysis of the key secondary endpoint 21
2.5 Determination of sample size 21
2.5.1 Sample size calculation for the primary endpoint analysis 21
2.5.2 Power for analysis of key secondary endpoint 22
2.6 Randomization and stratification 22
2.7 Definitions of analysis sets 23
2.7.1 Full Analysis Set (FAS) 23
2.7.2 Per-protocol Set (PPS) 24
2.7.3 Safety set (SAF) 24
3 Changes to planned analyses 25
4 Statistical analyses 25
4.1 General considerations 25
4.1.1 Baseline definition 25
4.1.2 Study Period definitions 25
4.1.3 Study day 26
4.1.4 Visit windows 26
4.1.5 Duration 27
4.2 Disposition of patients and protocol deviations 27
4.3 Demographics and baseline characteristics 27
4.3.1 Demographic and baseline disease characteristics 27

[redacted]

4.3.2 Medical history 28
4.3.3 Prior and concomitant medications 29
4.4 Treatment regimen, compliance and exposure to drug 29
4.4.1 Extent of exposure 29
4.4.2 Compliance 30
4.5 Analysis of the primary endpoint 31
4.6 Analysis of secondary endpoints 32
4.6.1 Efficacy evaluation 32
4.6.2 Safety evaluation 34
4.6.3 Immunogenicity 36
4.6.4 Pharmacokinetic evaluations 36
4.7 Handing of missing data in the analysis 36
5 Clinical Study Report – Appendix 16.1.9 Documentation of statistical methods 37
6 Appendices 38
6.1 Appendix 1: Adverse event partial date imputation 38
6.1.1 Adverse event end dates 38
6.1.2 Adverse event start date 38
6.2 Appendix 2: Concomitant medication partial date imputation 39
6.2.1 Concomitant medication end date 39
6.2.2 Concomitant medication start date 39
6.3 Appendix 3: Adverse events of special interest 40
6.4 Appendix 4: Medications of special interest 42
References 43

List of tables

Table 2-1 Objectives and related endpoints 9
Table 2-2 EULAR response criteria 13
Table 2-3 HAQ-DI items and aids 15
Table 2-4 FACIT Fatigue scale scoring 16
Table 2-5 Blood collection schedule for ADA and PK samples 19
Table 4-1 MMRM Analysis SAS® Code 31
Table 4-2 ANCOVA analysis SAS® Code 33
Table 6-1 ATC codes and preferred terms for DMARDs 42

List of figures

Figure 2-1 28-Joint count 12
Figure 2-2 Power of AUEC analysis for various common standard deviations 22

[redacted]

Abbreviations and definitions

ACR	American College of Rheumatology
ADA	Anti-Drug Antibody
AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BDRM	Blind Data Review Meeting
BMI	Body Mass Index
CI	Confidence Interval
CDISC	Clinical Data Interchange Standards Consortium
CM	Concomitant Medication
CRP	C-Reactive Protein
CSR	Clinical Study Report
DAS	Disease Activity Score
DMARD	Disease Modifying Anti-Rheumatic Drug
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
ESR	Erythrocyte Sedimentation Rate
EULAR	European League against Rheumatism
FAS	Full Analysis Set
FACIT	Functional Assessment of Chronic Illness Therapy
FUP	Follow-up
GA	Global assessment
HAQ	Health Assessment Questionnaire
IRT	Interactive Response Technology
ISR	Injection Site Reaction
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed-Model Repeated Measures
MTX	Methotrexate
PhGA	Physician’s global assessment of disease activity

[redacted]

PK	Pharmacokinetics
PPS	Per-Protocol Set
PT	Preferred Term
PtGA	Patient's global assessment of disease activity
RA	Rheumatoid Arthritis
RBC	Red Blood cell Count
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SCR	Screening
SDTM	Study Data Tabulation Model
SD	Standard Deviation
SJC66	Swollen Joint count on 66 Joints
SJC28	Swollen Joint Count on 28 Joints
SOC	System Organ Class
SP1	Study Period 1
SP2	Study Period 2
TEAE	Treatment Emergent Adverse Event
TJC68	Tender Joint Count on 68 Joints
TJC28	Tender Joint Count on 28 Joints
TNF	Tumor Necrosis Factor
TP	Treatment Period
VAS	Visual Analog Scale
WBC	White Blood cell Count
WHO	World Health Organization

[redacted]

1 Introduction

This document describes the detailed statistical methodology for the analysis of the Clinical Study GP17-302. It describes the planned analysis for Study Period 1 (SP1) of the study, i.e. up to the Week 24 visit, which includes the primary analysis at the Week 12 visit, and the analysis of Study Period 2 (SP2) data from Week 24 to Week 48, as well as the overall analysis of entire study data. This SAP covers all aspects of the statistical methodology with regard to the analysis of efficacy, safety, immunogenicity and local tolerability at the injection sites.

Any deviation from the final statistical analysis plan (SAP) after CDBL will lead to an SAP addendum and will be described in detail in the CSR, along with reasons for the change.

This version of the SAP will be finalized and approved in the Document Management System before the database is locked and before unblinding of the sponsor team.

Study Design

This study has a parallel-group, randomized, double-blind design and it is planned to randomize approximately 308 patients with moderate to severe active RA at approximately 110 study sites worldwide. A screening period of up to 4 weeks was used to assess patient´s eligibility. All patients had to be on a stable dose of MTX at least 4 weeks prior to randomization and continue throughout the study on this dose.

At the baseline visit (Day 1, Visit 2); eligible patients were randomized using a 1:1 ratio to one of the two treatment arms: GP2017 or US-licensed Humira®. Randomized patients entered a 24-week treatment phase Study Period 1 and were treated with s.c. injection of either GP2017 or US-licensed Humira®. Study drug administration was performed at the site by designated unblinded personnel. At Week 24 all patients were evaluated for their DAS28-CRP response (based on CRP taken at Week 22). Patients who did not show at least a moderate DAS28-CRP response had their final assessment at Week 24 and did not continue in the study.

Irrespective of their treatment during Study Period 1, all patients with at least a moderate response at Week 24 by DAS28-CRP score, were to be switched to or continued with subcutaneous injection of GP2017 from Week 24 up to Week 46. End of Study (EoS) Visit was to be performed at Week 48.

2 Statistical and analytical plans

Data will be analyzed by Sandoz according to the data analysis described in Section 9 of the study protocol. Important details are given in the following sections, and the statistical methods will also be summarized as applicable in Appendix 16.1.9 of the CSR.

2.1 Study documents and general considerations

2.1.1 Study documents

The following study documents are considered for the finalization of the SAP.

- Study protocol version 4, dated 01-Feb-2017
- Electronic Case Report Form (eCRF) version 8.0, dated 25-Sep-2017

[redacted]

- BDRM minutes version final 1.0, dated 19-Mar-2018

2.1.2 General considerations

Raw data listings, summary tables, figures and statistical tests will be generated using the SAS® Version 9.4 or higher.

All clinical data, including laboratory, immunogenicity, and pharmacokinetics data will be provided as raw data from an external database. CDISC SDTM 3.1.2 amendment 1 compliant SAS® datasets will be prepared. The final SDTM files will also include variables indicating the actual and planned treatment assigned.

Coding of corresponding data (e.g. by Medical Dictionary for Regulatory Activities [MedDRA] or World Health Organization [WHO]-drug dictionary), and ATC coding is included in the SDTM datasets.

Appropriate SAS® programs will be prepared and validated according to standard operating procedures.

The following descriptive statistics will be shown in summary tables:

- Continuous variables: n (number of patients with evaluable data), mean, standard deviation (SD), minimum, median, maximum. Quartiles will be presented as appropriate.
- Categorical variables: Count and percentage of each category. A row denoted “Missing” will be included in count tabulations where necessary to account for dropouts and missing values. Unless otherwise specified, the denominator for all percentages will be the number of patients in the treatment group within the analysis set of interest. Percentages will not be provided if the number is zero.

The same number of decimal places as in the raw data will be presented when reporting minimum and maximum, 1 more decimal place than in the raw data will be presented when reporting mean, median and quartiles, and 2 more decimal places than in the raw data will be presented when reporting SD and SE.

For the derived variables, appropriate precision should be provided in CSR appendix data listing. E.g., for the DAS28-CRP and DAS28-ESR, 2 decimal places will be kept in the data listing, the decimal places of the descriptive statistics will then follow the rules above accordingly (i.e., treat raw data as 2 decimal places).

In general, for by-visit summaries, data recorded at the nominal visit will be presented. With the exclusion of AEs (see in Section 2.3.2.1), only data up to study discontinuation (i.e., including End of Study visit) will be analyzed, but all the data will be listed. End of Study visit within the nominal visit windows on a condition of absence of a nominal visit will be assigned to the nominal visit (see in Section 4.1.4), otherwise it will be included as a separate visit regardless of the day it took place. Unless otherwise noted, unscheduled visits will not be included in by-visit summaries. However unscheduled measurements will be presented in the listings.

Date imputations will be performed as described in Appendix 1 and Appendix 2. Imputed dates will be used for further derivations, only the not imputed partial dates will be presented in the listings.

[redacted]

Study Period 1 summary tables will be presented by treatment as

- GP2017
- Humira
- Total

Study Period 2 summary tables will be presented by treatment as

- Continued GP2017
- Humira to GP2017
- Total

Overall (including entire study) summary tables will be presented by treatment as

- GP2017
- Humira/Switched GP2017
- Total

2.2 Study objectives and endpoints

The study objectives and related endpoints see Table 2-1.

Table 2-1 Objectives and related endpoints

Objectives	Endpoints
Primary objectives • To demonstrate similar efficacy of GP2017 and US-licensed Humira® in patients with moderate to severe active RA up to Week 12	**Primary endpoint** • DAS28-CRP score change from baseline up to Week 12
Key secondary objective • To demonstrate similar efficacy of GP2017 and US-licensed Humira® in patients with moderate to severe active RA until Week 24	**Key secondary endpoint** • Time-weighted averaged change from baseline in DAS28-CRP until Week 24
Other secondary objectives in Study Period 1 • To demonstrate similar efficacy and to compare safety and immunogenicity of GP2017 and US-licensed Humira® in patients with moderate to severe active RA over 24 weeks of treatment	**Other secondary endpoints in Study Period 1** • Proportion of GP2017 and US-licensed Humira® treated patients achieving EULAR criterion for remission at Week 4, 12, 24 • Proportion of GP2017 and US-licensed Humira® treated patients achieving EULAR criterion for good response at Week 4, 12, 24 • Proportion of GP2017 and US-licensed Humira® treated patients achieved EULAR moderate response at Week 4, 12, 24 • DAS28-CRP and DAS28-ESR changes from baseline in GP2017 and US-licensed Humira® treated patients at Week 2, 4, 24

[redacted]

Objectives	Endpoints
	• Proportion of GP2017 and US-licensed Humira® treated patients achieving remission according to Boolean definition (TJC (28 joints) ≤1, SJC (28 joints) ≤1, CRP≤1 mg/dl and patient global assessment (PtGA)≤1 on a scale of 1-10 (corresponding to ≤10 on a scale of 1-100)) at Week 4, 12, 24 • Proportion of GP2017 and US-licensed Humira® treated patients achieving ACR20/50/70 (based on CRP & ESR) responses at Week 4, 12, 24 • Health Assessment Questionnaire-Disability Index (HAQ-DI) changes from baseline in GP2017 and US-licensed Humira® treated patients at Week 4, 12 and 24 • Proportion of GP2017 and US-licensed Humira® treated patients achieving HAQ-DI score ≤0.5 at Week 4, 12 and 24 • Proportion of GP2017 and US-licensed Humira® treated patients with HAQDI score improvement >0.3 at Week 4, 12 and 24 • FACIT Fatigue scale changes from baseline in GP2017 and US-licensed Humira® treated patients at Week 4, 12 and 24 • CRP and ESR changes from baseline in GP2017 and US-licensed Humira® treated patients over time • Clinical safety of GP2017 and US-licensed Humira® as assessed by changes in vital signs, clinical laboratory parameters and incidence and severity of Adverse Events (AE) • Incidence and severity of injection site reactions (ISRs) in GP2017 and US-licensed Humira® treated patients • Immunogenicity as determined by measuring the rate of anti-drug antibody (ADA) formation against adalimumab in GP2017 and US-licensed Humira® treated patients
Other secondary objectives in Study Period 2 • To evaluate long-term safety, immunogenicity and efficacy of GP2017 up to Week 48 and to investigate the effects of a switch from US-licensed Humira® to the proposed biosimilar GP2017 in patients with at least a moderate response, with respect to efficacy, safety and immunogenicity	**Other secondary endpoints in Study Period 2** • Clinical safety in patients treated continuously with GP2017 and in patients treated with GP2017 after switch from US-licensed Humira® as assessed by changes in vital signs, clinical laboratory parameters and incidence and severity of AE • Incidence and severity of injection site reactions (ISRs) in patients treated continuously with GP2017 and in patients treated with GP2017 after switch from US-licensed Humira® • Immunogenicity as determined by measuring the incidence of ADA formation against adalimumab in patients treated continuously with GP2017 and in patients treated with GP2017 after switch from US-licensed Humira®

[redacted]

Objectives	Endpoints
	• Efficacy in patients treated continuously with GP2017 and in patients treated with GP2017 after switch from US-licensed Humira® as assessed by DAS28-CRP, DAS28-ESR score changes from Week 24 at Week 48 and ACR20/50/70 responses at Week 48 • Health Assessment Questionnaire-Disability Index (HAQ-DI) changes from Week 24 at Week 48 in patients treated continuously with GP2017 and in patients treated with GP2017 after switch from US-licensed Humira® • Proportion of patients treated continuously with GP2017 and patients treated with GP2017 after switch from US-licensed Humira® achieving HAQ-DI score ≤0.5 at Week 48 • Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue scale changes from Week 24 at Week 48 in patients treated continuously with GP2017 and in patients treated with GP2017 after switch from US-licensed Humira®
[redacted]	

2.3 Assessments

2.3.1 Efficacy assessments

2.3.1.1 Joint counts (TJC28, SJC28, TJC68 and SJC66)

A joint which is missing consistently throughout the study will be imputed as “not swollen” and “not tender”. This is the most commonly (76%) used procedure for imputation in publications of RA trials (Ibrahim et al, 2016).

Most joints will not have swelling or tenderness and hence a non-swelling, non-tender is the most plausible value to use for imputation. In most cases the joint cannot be assessed for a specific reason, e.g. surgery or joint replacement in which cases it is clinically appropriate to consider these joints as not having tenderness or swelling. Since we are primarily interested in the change in the DAS28 from baseline then the exact method of imputation will not have a substantial impact. With this in mind it is important that a joint has to be missing consistently throughout the study when this imputation method is applied. So for example, a joint replacement or surgery conducted during the trial would be considered a sufficient reason to exclude the patient from the PPS, although the patient would still be included within the FAS analysis.

The effect of non-swelling, non-tender imputation is that the DAS28 will be calculated with the lowest possible score. Hence, it is assured that patients that meet the inclusion criterion of

[redacted]

DAS28-CRP ≥ 3.2 at screening using this method of imputation would have met the inclusion criterion regardless of the value of the unassessed joints.

However, for this analysis to be meaningful it is recognized that there must be a minimum number of joints assessed beyond which the joint count will become unreliable. Hence it was decided that if more than two of the joints are missing then the 28-joint counts (TJC28 and SJC28) cannot be calculated. The full tenderness joint counts TJC68 will not be calculated if 4 or more joint counts are missing or if the TJC28 cannot be calculated. The full swelling joint counts SJC66 will not be calculated if 4 or more joint counts are missing or if the SJC28 cannot be calculated.

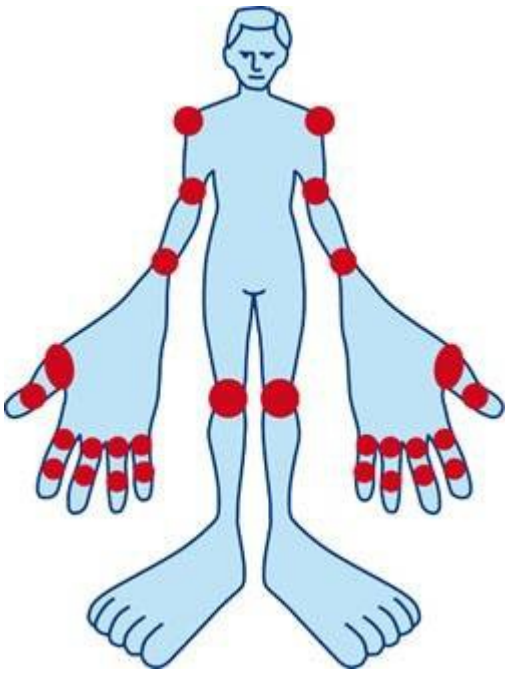
The number of missing joints at baseline will be presented in the baseline table and also be listed.

The following joints (see Figure 2-1) are included into the 28-joint count:

- Shoulder (left and right) – 2 joints
- Elbow (left and right) – 2 joints
- Wrist (left and right) – 2 joints
- MCP I to V (left and right) – 10 joints
- PIP I to V (left and right) – i.e. including thumbs – 10 joints
- Knee (left and right) – 2 joints

Note that distal inter-phalangeal joints are not included.

Figure 2-1 28-Joint count



[redacted]

2.3.1.2 DAS28-CRP

The disease activity score (DAS) is a combined index to measure the disease activity in patients with RA. The DAS28-CRP is calculated as

$$\text{DAS28-CRP} = 0.56 \times \sqrt{(\text{TJC28})} + 0.28 \times \sqrt{(\text{SJC28})} + 0.36 \times \ln(\text{CRP}+1) + 0.014 \times \text{PtGA} + 0.96$$

With TJC28, SJC28 measured as described in Section 2.3.1.1, the C-reactive protein (CRP) measured in mg/L units, and the patient's global assessment of disease activity (PtGA) measured in mm indicated by marking a 100 mm line between very good and very bad. For CRP values below the limit of detection (<1 mg/L) value of zero will be used in the DAS28-CRP calculation.

The DAS28-CRP provides a number on a scale from 0.96 to 10 indicating the current activity of the RA of the patient.

TJC28 and SJC28 for patients with partially missing joint counts can be imputed using imputation defined in Section 2.3.1.1. DAS28-CRP will not be imputed if any of the components (TJC28, SJC28, CRP or PtGA) are missing.

Due to several days needed to obtain the CRP results from central laboratory, the decision to continue patients in the study after Week 24 was to be made based on Week 24 DAS28 parameters and the CRP value from Week 22. However for the statistical analysis the DAS28-CRP score at Week 24 will be calculated based on the Week 24 CRP value provided by the central laboratory. The analysis date for DAS28-CRP will be the date of the nominal visit.

2.3.1.3 DAS28-ESR

The DAS28-ESR is calculated as

$$\text{DAS28-ESR} = 0.56 \times \sqrt{(\text{TJC28})} + 0.28 \times \sqrt{(\text{SJC28})} + 0.70 \times \ln(\text{ESR}) + 0.014 \times \text{PtGA}$$

With TJC28, SJC28 measured as described in Section 2.3.1.1, the ESR measured using mm/h units and PtGA measured in mm. ESR in adults is expected to be ≥ 1 mm/h, however in some cases the lab kit can record a zero value. In these cases the value of 1 mm/h will be used in the DAS28-ESR calculations above to make the formula calculable. The analysis date for DAS28-ESR will be the date of the nominal visit.

The DAS28-ESR provides a number on a scale from 0 to 10 indicating the current activity of the RA of the patient. The DAS28-ESR will be missing if any of the TJC28, SJC28, ESR or PtGA are missing.

2.3.1.4 EULAR response criteria

The EULAR good response, moderate response, and no response based on the DAS28-CRP are defined as follows.

Table 2-2 EULAR response criteria

	Improvement in DAS28 from baseline		
Present DAS28	> 1.2	> 0.6 and ≤ 1.2	≤ 0.6
≤ 3.2	good response	moderate response	no response

[redacted]

	Improvement in DAS28 from baseline		
Present DAS28	> 1.2	> 0.6 and ≤ 1.2	≤ 0.6
> 3.2 to ≤ 5.1	moderate response	moderate response	no response
> 5.1 (high)	moderate response	no response	no response

EULAR remission is defined as DAS28-CRP < 2.6.

If the DAS28-CRP is missing then the EULAR response is set to missing. Note that it is a condition of continuation into SP2 that a patient show at least a moderate response based on DAS28-CRP at Week 24 (using Week 22 CRP values).

2.3.1.5 American college of rheumatology (ACR) response criteria

This efficacy variable is the clinical response to treatment according to ACR improvement criteria. A patient will be considered a responder according to ACR20 criteria if she/he fulfills all following three criteria:

- at least 20% improvement from baseline in tender joint count, using the 68-joint count
- at least 20% improvement from baseline in swollen joint count, using the 66-joint count
- and at least 20% improvement from baseline in at least 3 of the following 5 measures:
  - Patient's assessment of RA pain (VAS 100 mm)
  - Patient's global assessment of disease activity (VAS 100 mm)
  - Physician's global assessment of disease activity (VAS 100 mm)
  - Patient self-assessed disability (Health Assessment Questionnaire [HAQ] Disability Index)
  - CRP or ESR

ACR50 and ACR70 are defined as ACR20 replacing '20% improvement' by '50% improvement' and '70% improvement' respectively.

The ACR criteria will be calculated for Week 4, Week 12, Week 24 and Week 48 (Visits 1, 2, 4, 8, 14 and 16).

2.3.1.6 Boolean remission

Boolean remission is defined as TJC28 ≤ 1 and SJC28 ≤ 1, CRP level (mg/dL) ≤ 1, and PtGA (VAS 100 mm) ≤ 10.

If any of the TJC28, SJC28, CRP, or PtGA are missing then the Boolean remission is missing.

2.3.1.7 Patient's assessment of RA pain

The patient's assessment of pain was to be assessed by the patient as defined in the protocol at screening, baseline, Week 2, Week 4, Week 12, Week 24 and Week 48 (Visits 1, 2, 3, 4, 8, 14 and 16) using 100 mm visual analog scale (VAS) ranging from 'no pain' to 'unbearable pain'.

[redacted]

2.3.1.8 Patient's global assessment of disease (PtGA)

The patient's global assessment of disease activity will be assessed by the patient as defined in the protocol at screening, baseline, Week 2, Week 4, Week 12, Week 24 and Week 48 (Visits 1, 2, 3, 4, 8, 14 and 16) using 100 mm VAS ranging from "very good" to "very poor"..

2.3.1.9 Physician's global assessment of disease activity (PhGA)

The physician's global assessment of disease activity will be assessed by the investigator as defined in the protocol at screening, baseline, Week 2, Week 4, Week 12, Week 24 and Week 48 (Visits 1, 2, 3, 4, 8, 14 and 16) using 100 mm VAS ranging from "very good" to "very poor",

2.3.1.10 Health-related quality of life disability index (HAQ-DI)

The patient health assessment questionnaire disability index, HAQ-DI, will be used to assess physical ability and functional status of patients as well as quality of life.

HAQ-DI assessments will be done at Baseline, weeks 4, 12, 24, and 48 or, if applicable, at End of Study Visit.

The HAQ-DI consists of 20 questions, or items, which are grouped into 8 categories, or sections, of functioning in performing 8 common activities of daily living (with 2 or 3 questions in each category). Scoring within each section is from 0 (without any difficulty) to 3 (unable to do). For each section the score given to that section is the worst score within the section, i.e. if one question is scored 1 and another 2, then the score for the section is 2. In addition, if an aide or device is used or if help is required from another individual, then the minimum score for that section is 2. If the section score is already 2 or more then no modification is made.

Table 2-3 HAQ-DI items and aids

Section	Daily activity	Items	Aids and help
1	Dressing and grooming	(1) able to dress yourself (2) able to use shampoo	Devices used for dressing, Help dressing or grooming
2	Arising	(1) able to stand up from straight chair (2) able to get out of bed	Special or built up chair, help arising
3	Eating	(1) able to cut your meat (2) able to lift a full cup or glass to your mouth (3) able to open a new milk carton	Built up or special utensils, help eating
4	Walking	(1) able to walk outdoors on flat ground (2) able to climb up five steps	Cane, Walker, Crutches, Wheelchair, help walking
5	Reaching	(1) able to reach down to get a 5 pound object from just above your head (2) able to bend down to pick up clothing from the floor	Long-handled appliances for reach, help reaching

[redacted]

Section	Daily activity	Items	Aids and help
6	Hygiene	(1) able to wash and dry your body (2) able to take a tub bath (3) able to get on and off the toilet	Bathtub bar, Bathtub seat, Long-handled appliances in bathroom, Raised toilet seat, help with hygiene
7	Gripping	(1) able to open car doors (2) able to open jam jars which have been previously opened (3) able to turn faucets on and off	Jar opener, help with gripping and opening things
8	Activities	(1) able to run errands and shop (2) able to get in and out of a car (3) able to do chores such as vacuuming or yardwork	Help with errands and chores

The 8 scores of the 8 sections are summed and divided by 8 to provide the total HAQ-DI score. The range for this score is (0, 3). In the event that one/two sections are not completed by a patient then the summed score would be divided by 7/6. The HAQ-DI total score is not computed when the patient provides answers in fewer than 6 categories.

2.3.1.11 Functional assessment of chronic illness therapy - fatigue scale (FACIT- Fatigue)

The Functional Assessment of Chronic Illness Therapy – Fatigue Scale (FACIT-Fatigue) is a 13- item questionnaire that assesses self-reported fatigue and its impact upon daily activities and function. FACIT assessments will be completed by the patient at Baseline, weeks 4, 12, 24, and 48 or, if applicable, at End of Study Visit.

Table 2-4 FACIT Fatigue scale scoring

No.	Code	Question	Not at all	A little bit	Some-what	Quite a bit	Very much
1	HI7	I feel fatigued	4	3	2	1	0
2	HI12	I feel weak all over	4	3	2	1	0
3	An1	I feel listless (“washed out”)	4	3	2	1	0
4	An2	I feel tired	4	3	2	1	0
5	An3	I have trouble starting things because I am tired	4	3	2	1	0
6	An4	I have trouble finishing things because I am tired	4	3	2	1	0
7	An5	I have energy	0	1	2	3	4
8	An7	I am able to do my usual activities	0	1	2	3	4
9	An8	I need to sleep during the day	4	3	2	1	0
10	An12	I am too tired to eat	4	3	2	1	0

[redacted]

No.	Code	Question	Not at all	A little bit	Some-what	Quite a bit	Very much
11	An14	I need help doing my usual activities	4	3	2	1	0
12	An15	I am frustrated by being too tired to do the things I want to do	4	3	2	1	0
13	An16	I have to limit my social activities because I am tired	4	3	2	1	0

Items are scored as follows: 4=Not at All; 3=A little Bit; 2=Somewhat; 1=Quite a bit; 0=Very much, except for items number 7 and 8 which are reverse scored.

Each of the 13 items of the FACIT-Fatigue scale ranges from 0-4 (0 is the worst response), and are equally scaled so that the range of possible FACIT Fatigue scores is 0-52, with 0 being the worst possible score and 52 the best.

In cases where some answers may be missing, a total score is prorated from the score of the answered items, so long as more than 50% of the items (i.e., at least 7 of 13) were answered.

2.3.1.12 Rheumatoid factor and anti-CCP antibodies

Rheumatoid factor and anti-CCP antibodies assessment as a measure of RA serological status will be performed at screening, week 12, week 24, week 48 / End of Study visit. Rheumatoid factor values ≤ 10 UI/mL and anti-CCP antibodies values < 17 U/mL are considered negative, otherwise positive.

2.3.2 Safety and tolerability assessments

2.3.2.1 AEs

An adverse event is the appearance or worsening of any undesirable sign, symptom, or medical condition occurring after starting the study drug even if the event is not considered to be related to study drug. Study drug includes the investigational drug under evaluation and the comparator drug that is given during any phase of the study. Medical conditions/diseases present before starting study drug are only considered adverse events if they worsen after starting study drug. Abnormal laboratory values or test results constitute adverse events only if they induce clinical signs or symptoms, are considered clinically significant, or require therapy.

Adverse events will be coded using MedDRA version 20.0.

Treatment emergent AEs (TEAEs) are events started after the first dose of study treatment and before study discontinuation or 30 days after last dose, whichever occurs later. In case where it is not possible to define an AE as TEAE or not, the AE will be classified by the worst case, i.e. treatment emergent.

All adverse events with start date on or after the date of the first study medication in SP2 will be counted as belonging to Study Period 2. In the case of a partial start date see Appendix 1 on how such cases will be handled. If it is unclear whether the AE belongs to SP1 or SP2 then the AE will be assigned to SP2. This is a conservative approach since all patients are treated with the investigational product GP2017 in SP2.

[redacted]

Adverse events of special interest are defined as a subset of TEAE in Appendix 3.

The investigator or qualified designee will have recorded ISRs belonging to one of the following: injection site erythema, injection site haematoma, injection site induration, injection site inflammation, injection site nodule, injection site oedema, injection site pain, injection site pruritus, injection site rash or injection site discoloration at the specified time points. Incidence and occurrence of ISRs and symptoms will be calculated.

All ISR are considered as a subgroup of TEAE. An ISR that fulfills the criteria of an SAE will be documented and reported as such.

2.3.2.2 Laboratory parameters

For the identification of notable values, the laboratory manual should be consulted. The following laboratory categories will be listed.

Hematology

Blood will have been obtained at screening, baseline, week 4, week 12, week 24, week 36 and week 48 (or as required at an unscheduled visit) to measure erythrocytes, hematocrit, hemoglobin, leukocytes, platelets, basophils (absolute and percentage), eosinophils (absolute and percentage), lymphocytes (absolute and percentage), monocytes (absolute and percentage), neutrophils (absolute and percentage).

Clinical Chemistry

Blood will be obtained at screening, baseline, week 4, week 12, week 24, week 36 and week 48 (or as required at an unscheduled visit) to measure alanine aminotransferase, albumin, alkaline phosphatase, aspartate aminotransferase, total bilirubin, calcium, serum creatinine, gamma glutamyl transferase, potassium, total protein, sodium, uric acid, total cholesterol, HDL cholesterol, LDL cholesterol, rheumatoid factor, cyclic citrullinated peptide antibody, immunoglobulin A, immunoglobulin G,

Coagulation

Activated Partial Thromboplastic Time (aPTT) and INR.

Urinalysis

Dipsticks will be provided by the central laboratory to the sites for local urinalysis assessments. The sites will record the results in the appropriate eCRF page for each patient.

Measurements for density, pH, glucose, cells (RBC, WBC), proteins, blood and cylinders will be done at screening, baseline, week 4, week 12, week 24, week 36 and week 48 (or as required at an unscheduled visit) .

Vital signs

Vital signs include blood pressure and heart rate measurements and will be assessed at all visits. The repeat sitting measurements will be made at 1-2 minute intervals and the mean of the three measurements will be used and entered on the corresponding eCRF page.

[redacted]

ECG

Standard 12-lead ECG measurements will be assessed at time points as per the study schedule (screening, week 24 and week 48). ECG results are captured in the eCRF as "Normal" "Abnormal, clinically insignificant" and "Abnormal, clinically significant". Any abnormal clinically significant ECG findings should be recorded as adverse events.

Physical examination

A complete physical examination will include the examination of general appearance, skin, neck (including thyroid), eyes, ears, nose, throat, lungs, heart, abdomen, back, lymph nodes and extremities, as well as vascular and neurologic examination and will be done at screening, baseline, week 24 and week 48. A short physical examination with assessment of general appearance will be done at all other visits.

2.3.2.3 Immunogenicity

Immunogenicity of Adalimumab as determined by the formation of anti-drug-antibodies (ADAs) will be assessed by using validated immunoassays. Immunogenicity data will be received as external source data for analysis.

Blood samples for immunogenicity (assessment of anti-adalimumab antibodies development) will be taken at specified time points as indicated in Table 2-5

Table 2-5 Blood collection schedule for ADA and PK samples

Study visits	Time/Week
Visit 2	Baseline
Visit 3	Week 2
Visit 4	Week 4
Visit 8	Week 12
Visit 14	Week 24
Visit 15	Week 36
Visit 16	Week 48
End of Study Visit	If applicable
Unscheduled Visit	If applicable

All samples were analyzed in a screening assay. Study samples with a result below the validated screening cut-point were negative for binding anti-adalimumab antibodies and will be reported accordingly. In the event of a positive result (result above the screening cut-point) the sample were additionally analyzed in a secondary confirmatory assay (specificity assay). In case the assay signal could be reduced after addition of excess of adalimumab beyond the validated confirmatory cut-point, a sample was reported as confirmed binding positive. In contrast, samples with a result above the screening cut-point in the screening assay but which were negative in the confirmatory assay were reported as negative.

A transient ADA response occurs where a post baseline positive ADA response occurs but is followed by a negative ADA response before the end of Treatment Period 1or end of study visit.

The incidence of patients with neutralizing ADAs will be summarized by treatment arms.

[redacted]

2.3.2.4 Pharmacokinetic evaluations

No formal PK will be performed. The adalimumab trough serum levels will be assessed at the same time points as ADA to support the evaluation of immunogenicity. Thereby, the drug tolerance level of the ADA method (which was determined during method validation) can be compared with the determined drug levels.

2.4 Statistical method planned in the protocol

The following section outlines the statistical methods planned in the protocol.

2.4.1 Analysis of the primary endpoint

The following statistical hypotheses will be used to assess equivalence between GP2017 and US-licensed Humira® with regards to DAS28-CRP change from baseline at Week 12:

H0: |GP2017 – Humira®| ≥ 0.6 versus H1: |GP2017 – Humira®| < 0.6

To address both EMA and FDA regulatory requirements for establishing therapeutic equivalence both 95% and 90% confidence intervals will be estimated and used to assess equivalence of DAS28-CRP change from baseline at Week 12.

Therapeutic equivalence in terms of DAS28-CRP will be concluded if the 95% (as required by EMA) or 90% (as required by FDA) confidence interval for the difference in DAS28-CRP change from baseline at Week 12 between the two treatment groups is completely contained within the interval [-0.6, 0.6]. This is statistically equivalent to calculating two independent one-sided tests at 2.5%/5% alpha level (one in each direction), of which both have to be successful.

A mixed-model repeated measures (MMRM) analysis will be performed for DAS28-CRP change from baseline up to Week 12 as the endpoint, including treatment, stratification factors (region, body weight category, prior systemic therapy), time, the interaction between time (visits) and treatment all as categorical variables, and baseline DAS28-CRP as a continuous variable in the model.

DAS28CRP CFB = treat + strata + time + (time*treat) + baseline DAS28CRP

In the case that the IRT and eCRF values of the stratification factors differ the correct definition of the stratification factors used for the primary analyses will be based on the clinical database.

MMRM is a standard approach to longitudinal analysis of continuous endpoints. Its roots bear into linear mixed modeling methodology and it is specified as a multivariate normal model of the longitudinal data. More specifically, this model includes a saturated visit-by-treatment structure for the mean. That is, time is considered as a factor rather than a continuous variable, and thus MMRM makes no assumption about trends over time. Traditional linear mixed models make further assumptions about the covariance of within/intra-patient observations, which can either originate from random effects or residual errors. With MMRM the covariance structure is specified only through the residual error term. The most flexible covariance matrix is unstructured, that is, all variance and covariance parameters are estimated. Therefore, an unstructured covariance matrix will be used for the analysis, and if the unstructured covariance matrix results in a lack of convergence, an appropriate covariance structure will be selected.

[redacted]

Mean change from baseline at Week 12, standard errors and the two-sided 95%/90% CI for the mean difference between GP2017 and US-licensed Humira® will be estimated from the model and the 95%/90% CI compared to the pre-specified equivalence margin of [-0.6; 0.6]. A 0.6 change in DAS28-CRP score is considered as no clinically meaningful difference by EULAR criteria (Table 2-2) and is therefore used as the equivalence margin limits [0.6,-0.6].

The primary analysis will be performed on the Week 12 Per-protocol set (W12 PPS), which is the most appropriate analysis set to use when testing for equivalence.

2.4.2 Analysis of the key secondary endpoint

Time-weighted averaged change from baseline in DAS28-CRP until Week 24 (standardized AUEC approach) will be calculated and analyzed using an ANCOVA.

The averaged change from Week 2 and Week 24 will be estimated by a time-weighted mean across visits. The weights are derived using the standard formula from the trapezoidal rule across the visits. Mean averaged change from baseline from Week 2 to Week 24, standard errors and the two-sided 95%/90% CI for the mean difference between GP2017 & and US-licensed Humira® will be estimated from the model and the 95%/90% CI compared to the pre-specified equivalence margin of [-0.6; 0.6].

Both 95% and 90% confidence intervals shall be estimated and used to assess equivalence of time-weighted averaged change from baseline in DAS28-CRP until Week 24. No imputation will be performed for missing components of the DAS28-CRP score or the computed AUEC value itself.

The key secondary analysis will be performed on the Study Period 1 Per-protocol set (SP1 PPS) which is the most appropriate analysis set to use when testing for equivalence.

2.5 Determination of sample size

2.5.1 Sample size calculation for the primary endpoint analysis

The sample size for this study is based on an expected difference of zero and common SD of 1.30 for the DAS28-CRP change from baseline at Week 12 (ReAct study), an equivalence margin of 0.6 and assumed loss of 20% patients from the per-protocol analysis set. A 0.6 change in DAS28-CRP score is considered as no clinically meaningful difference by EULAR criteria and is therefore used as the equivalence margin limits [-0.6; 0.6].

In order to fulfil FDA regulatory requirements for establishing therapeutic equivalence (e.g. using 90% confidence interval), a sample size of 154 patients per treatment group (to maintain overall 246 evaluable patients for the primary endpoint analysis) will have 95% power to test the equivalence between GP2017 and US-licensed Humira®.

In order to fulfil EMA regulatory requirements for establishing therapeutic equivalence (e.g. using 95% confidence interval), a sample size of 154 patients per treatment group (to maintain overall 246 evaluable patients for the primary endpoint analysis) will have 90% power to test the equivalence between GP2017 and US-licensed Humira®.

[redacted]

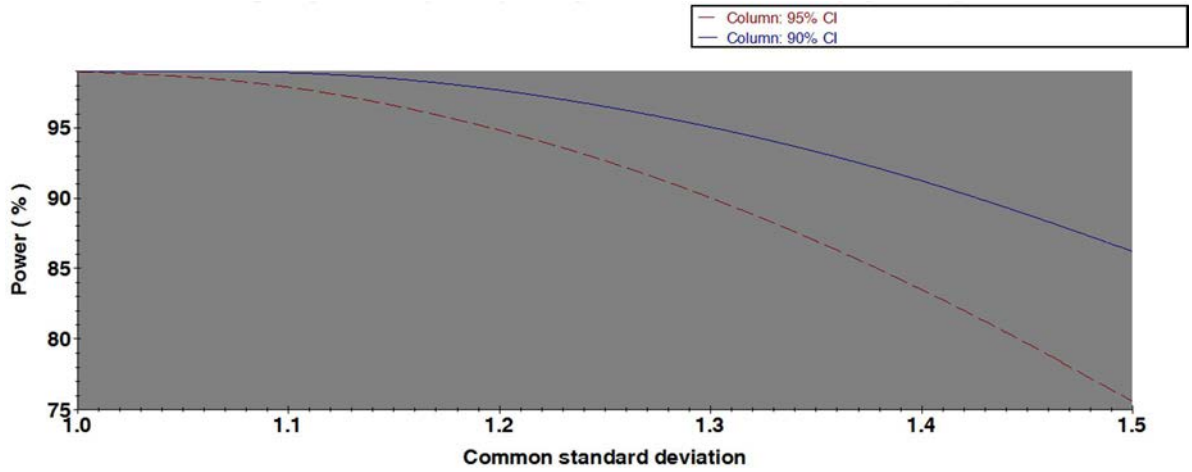
2.5.2 Power for analysis of key secondary endpoint

From Burmester et al (2014) the standard deviation of the change from baseline of the DAS28-CRP at 6 months, i.e. roughly Week 24, was observed as 1.5. This is comparable with the value of 1.3 found by Burmester et al (2014) for the Week 12 assessment. This matches well with Wells et al (2009) who found that the standard deviation of the DAS28-CRP was 0.9 to 1.0, which would mean a change from baseline measurement could be expected to have an SD of 1.4 times of the baseline SD (provided the baseline and post-baseline measurements were independent).

The time weight average is a weighted score of four change values where the weights sum to 1. The exact value of the SD for the time-weight average is unknown because it depends on both the days of each visit for each subject and the correlation between the individual measurements. However, a range for the SD can be estimated. Assuming each DAS28-CRP change from baseline value has a common SD of 1.4 then the lowest possible SD for the time-weighted average will occur if all four changes scores are independent and equally spaced, so that the weights are all 0.25, and this is then the same as taking an average of four independent random variables, and in this case the standard deviation of the time-weighted value will be ½ of common SD, i.e. 0.7. At the other extreme the maximum of the weighted average occurs if each change score is perfectly correlated with the others. Then the specific values of the weights are irrelevant as the weighted average, will be the same as a single value of the change score, and hence will have an SD of 1.4.

Figure 2-2 below shows the resulting power for different values of the SD for the time-weighted average with 246 evaluable patients when calculating both 90% and 95% confidence interval. It is clear that there is always sufficient power (at least 80%) to demonstrate equivalence using the 90% confidence interval, and that the power will also be sufficient using the 95% confidence interval.

Figure 2-2 Power of AUEC analysis for various common standard deviations



2.6 Randomization and stratification

Randomization will be stratified at baseline (Visit 2) by region, body weight and prior therapy. The stratification ensures balanced allocation of patients to treatment groups within the strata.

[redacted]

- The strata for body weight will be "body weight < 80 kg" or "body weight ≥ 80 kg". Stratification by body weight will be based on the weight assessment at Visit 2.
- The strata for region will be "RoW" and "Americas".
- The strata for prior therapy are presented in hierarchical order and the option highest in hierarchy has to be chosen in case of more than 1 prior therapy scheme in the individual patient's history.
    1. Prior treatment with biologic DMARDs (patient belongs to this strata even if treated with other synthetic DMARDs as well)
    2. Prior treatment with other conventional synthetic DMARDs") patient belongs to this strata only if they did not receive prior treatment with biologic DMARDs),
    3. Prior treatment with MTX only

Considering the fact that the treatment arm assigned by the IRT cannot be revised for a patient if a wrong stratum was selected, the correct value of the stratification factor coming from the clinical database will be used for the statistical analyses. However, summaries and listing of patients randomized in the incorrect stratum will also be presented (see Section 4.3.1.1).

The stratification factors will be included as factors in the statistical methods described in Section 2.4.

2.7 Definitions of analysis sets

The following analysis sets will be used for this analysis.

The primary endpoint analysis will be based on W12 PPS, and will be repeated on SP1 FAS as sensitivity analysis.

The key secondary endpoint analysis will be based on SP1 PPS, and will be repeated on SP1 FAS as sensitivity analysis.

The analyses of Study Period 1 data will be based on SP1 FAS, SP1 SAF, SP1 PPS.

The analyses of Study Period 2 data will be based on SP2 FAS, SP2 SAF, SP2 PPS.

In addition to the study protocol defined objectives, the entire study data (i.e., from the randomization up to Week 48) will also be analyzed based on SP1 FAS, SP1 SAF, SP2 PPS.

2.7.1 Full Analysis Set (FAS)

Study Period 1 Full analysis set (SP1 FAS): consists of all randomized subjects to whom IMP has been assigned. Following the intent-to-treat principle, subjects will be analyzed according to the treatment assigned at randomization. Subjects will be analyzed as per the actual true strata and not necessarily the assigned strata.

Study Period 2 Full analysis set (SP2 FAS): The SP2 FAS set includes all SP1 FAS patients entering Study Period 2. Patients will be analyzed according to the treatment group assigned in SP1 and per the actual true strata from the clinical database.

[redacted]

2.7.2 Per-protocol Set (PPS)

Week 12 Per-protocol analysis set (W12 PPS): consists of all patients in the SP1 FAS who complete Week 12/Visit 8 and do not have any major protocol deviations regarding the evaluation of the study's primary objective and have received at least 5 doses of IMP up to Week 10.

The major protocol deviations were defined in BDRM, refer to the BDRM minutes for the details (version final 1.0, dated 19-Mar-2018).

Study Period 1 Per-protocol set (SP1 PPS): consists of all patients in the W12 PPS who complete Week 24/Visit 14 and do not have any major protocol deviations regarding the evaluation of the study's key secondary objective and have received at least 5 doses of IMP from Week 12 to Week 22.

Study Period 2 Per-protocol set (SP2 PPS): consists of all patients in the SP1 PPS who complete the entire study (Week 48/Visit 16) and do not have any major protocol deviations during the entire study with regards to any efficacy assessment and have received at least 10 doses of IMP from Week 24 to Week 46.

As defined in the PPS definitions, in addition to major protocol deviations, patients who completed the required study visit but did not meet the study treatment compliance will also be excluded from the per-protocol sets:

- For W12 PPS, patients have to receive at least 5 doses of IMP up to Week 10. Patients should receive six doses of IMP from Baseline to Week 10 (i.e., Baseline, Week 2, 4, 6, 8, 10).A patient who completed Week 12 (i.e., Week 12 visit assessment is available in subject visit dataset) but missed 2 or more doses of IMP will be excluded from W12 PPS.
- For SP1 PPS, patients have to receive at least 5 doses of IMP from Week 12 to Week 22. Patients should receive six doses of IMP from Week 12 to Week 22 (i.e., Week 12, 14, 16, 18, 20, 22). A patient who completed Study Period 1 (based on disposition dataset) but missed 2 or more doses of IMP will be excluded from SP1 PPS.
- For SP2 PPS, patients have to receive at least 10 doses of IMP from Week 24 to Week 46. Patients should self-administrate twelve doses of IMP biweekly from Week 24 to Week 46 (i.e., the EPOCH variable equals to "Treatment Period 2" in study drug exposure dataset). A patient who completed the Study Period 2 (based on disposition dataset) but missed 3 or more doses of IMP will be excluded from SP2 PPS.

2.7.3 Safety set (SAF)

Study Period 1 Safety set (SP1 SAF): consists of all subjects who received at least one dose of IMP, whether randomized or not. Subjects will be analyzed according to the treatment received.

Study Period 2 Safety set (SP2 SAF): The SP2 SAF set includes all SP1 SAF patients entering Study Period 2 who received at least one dose of IMP in Study Period 2.

[redacted]

3 Changes to planned analyses

Based on the study protocol Section 2.2, the efficacy evaluation of Study Period 2 includes calculating the changes from Week 24 at Week 48 for the following variables: DAS28-CRP, DAS28-ESR, HAQ-DI score, and FACIT fatigue score. In addition to changes from Week 24 at Week 48, the changes from baseline (i.e., randomization) at all applicable Study Period 2 visits will also be calculated and analyzed for above variables.

In addition to the study protocol defined objectives for SP1 and SP2 separately, the entire study data (i.e., from the randomization up to Week 48) will also be analyzed.

The SAP section "Power for analysis of key secondary endpoint", more detailed information than in the study protocol are provided.

The protocol specifies no imputation for missing components of the DAS28-CRP score, because it was assumed that joint surgery would not occur. In accordance with general practice, the joints which not assessable throughout the study will be considered as not having tenderness or swelling (see Section 2.3.1.1).

4 Statistical analyses

4.1 General considerations

4.1.1 Baseline definition

Unless stated otherwise, baseline is defined as the pre-dose assessment taken at Visit 2. For FACIT, HAQ-DI and Tender/swollen joints assessments baseline is defined on the questionnaire level, not on an individual measurements level.

For quantitative measurements, change and percent change from baseline will be calculated as:

- Change from baseline = Post-baseline value – Baseline value
- Percent change from baseline = 100% * (Post-baseline value – Baseline value) / Baseline value for baseline values greater than 0.

Changes from baseline will only be summarized for patients having both baseline and post baseline values.

For the changes from Week 24 at Week 48 in terms of DAS28-CRP, DAS28-ESR, HAQ-DI score, and FACIT Fatigue scale, the nominal Week 24 and Week 48 visit values will be used.

4.1.2 Study Period definitions

Day 1 of Study Period 1, i.e. the day on which the first dose of study medication is taken, is considered as the start for the SP1 and the study. All data before are considered belonging to Screening (SCR) period.

A patient is considered to have entered SP2 only if they have taken study medication assigned as belonging to SP2. The start of SP2 is defined as the first date of administration of study medication as assigned in SP2.

[redacted]

All data after study discontinuation (for AEs only in case they are after 30 days post last dose, see in Section 2.3.2.1) will be considered belonging to Follow-up (FUP) period. Concomitant medications started on day of discontinuation will be considered belonging to FUP period.

In cases where the last non-missing assessment and the start date coincide, the assessment will in general be considered pre-treatment (baseline), however AEs and medications commencing on the date of the first study medication injection will be considered post-baseline (SP1). Similarly, assessment on the day of the start date of SP2 will be assigned to SP1, while medications and AEs commencing on the day of the start date of SP2 will be assigned to SP2.

In the case of a partial date where it is unclear whether adverse events belong to SP1 or SP2 then they will be assigned to SP2. This is the more conservative assessment as all patients in SP2 are receiving the investigational product GP2017. Medications where it is unclear if they started in SP1 or SP2, i.e. there is a partial start date and no stop date in SP1, will be assigned to SP1 and reported in both periods.

Rules to impute partial dates for adverse events and concomitant medications for the purpose of assigning them to the correct study period are provided in Appendix 1 and Appendix 2. Completely missing dates will not be imputed.

4.1.3 Study day

Study day will be calculated from the reference start date, and will be used to show start/stop day of assessments and events.

Reference start date is defined as the day of the first dose of study drug administration (Day 1) and will appear in every listing where an assessment date or event date appears.

If the date of the event is on or after the reference date then:
Study Day = (date of event - reference date) + 1

If the date of the event is prior to the reference date then:
Study Day = (date of event – reference date)

In the situation where the event date is partial or missing, Study Day, and any corresponding durations will appear missing in the listings.

4.1.4 Visit windows

Study data for a scheduled visit (weeks 2, 4, 12, 24 and 48) should be entered on to the correctly designated eCRF page for that visit. However, in order to be considered per protocol a visit needs to fall within a window of time around the target date for the visit. In the protocol this is defined as a window of ±2 days for SP1 and ±3 days for SP2 around the target date, calculated from baseline (first dose date of IMP). Assessments that fall outside of this window will not be considered as being per protocol. However, all such assessments will be included into analysis and listings.

For the patients who early discontinued from the study, the End of Study visit will be mapped to the corresponding scheduled visit based on the study day.

[redacted]

4.1.5 Duration

All duration variables in days will be calculated as end date – start date + 1.

To get duration in month duration in days will be divided by 30.4375; duration in years – by 365.25.

4.2 Disposition of patients and protocol deviations

Frequency and percentage of patient's recruitment by region, country, site, and treatment group will be presented.

Disposition of patients includes the number and percentage of patients for the following categories: randomized, completed Week 12, discontinued before Week 12, completed SP1, discontinued during SP1, eligible to enter SP2 (achieving at least a moderate EULAR response assessed by DAS28-CSR at Week 24), entered SP2, completed SP2, discontinued during SP2. The primary reason for discontinuation in SP1/SP2 will be summarized.

A patient who completed the Visit 14 (Week 24) is considered to have completed SP1.

Frequency and percentage of patients having major and minor protocol deviations will be presented by deviation category, treatment group and overall for all patients in FAS. Patient having major and minor protocol deviations will be counted only once under more severe deviation category. Patients who are excluded from each analysis population and protocol deviations will also be listed.

All percentages for the summaries will be based on the number of patients in FASs (SP1 FAS for study period 1 and during the entire study, or, SP2 FAS for study period 2).

4.3 Demographics and baseline characteristics

4.3.1 Demographic and baseline disease characteristics

Demographic characteristics include categorical data (i.e., sex, race, and ethnicity) and continuous data (i.e., age, height, body weight, and body mass index [BMI]) will be summarized.

Age (years) = INT ((date of ICF – date of birth)/365.25)

The randomization stratification factors including region, weight category as per CRF and IRT, and prior therapy as per CRF and IRT will be summarized.

RA disease history will include duration since initial diagnosis of rheumatoid arthritis (in years), current global functional status in RA according to ACR, number of previous DMARDs used, previous MTX duration (in months), baseline MTX dose, and number of not assessable joints on baseline (out of 28 and full joint counts).

When counting the number of not assessable joints (out of 28/68) at baseline, in case a joint is not assessed at baseline, but assessed on screening or post-baseline visits (i.e. there was no underlying medical reason such as surgery), the joint will be considered assessable.

Partial dates of initial diagnosis will be imputed by earliest possible date (i.e. first day of month and 1st of January). MTX and DMARD treatments will be assessed using ATC coding. When

[redacted]

counting the number of DMARDs (excluding MTX) only distinct medications should be counted.

Previous MTX duration will be calculated from the date of earliest recorded MTX intake (partial dates will be imputed according to Appendix 2) up to date of first dose of study treatment.

RA baseline disease characteristics will include baseline values of: DAS28-ESR, DAS28-CRP, number of tender/swollen joints (TJC68, SJC66, TJC28, SJC28), CRP (mg/dL), ESR (mm/h), rheumatoid factor (positive/negative), anti-CCP antibodies (positive/negative), and the three VAS assessments (physician, patient, pain).

Other baseline characteristics (e.g. FACIT scores, HAQ-DI scores, tuberculosis, hepatitis and pregnancy screening) will also be presented in a summary table.

Summaries for continuous and categorical variables (as described in Section 2.1.2) will be presented for each treatment group and in total. All data will be listed.

4.3.1.1 Incorrect Stratum assignment

A table will be produced for each stratification factor which shows the number of patients in each of the strata assigned by the IRT system compared to the number of patients who should have been included into each stratum according to the clinical database. All stratification factors will be listed.

4.3.2 Medical history

Relevant medical history/current medical conditions not related to rheumatoid arthritis will be summarized and listed by Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary (version 20.0) system organ classes (SOC) and preferred terms (PT) on SAF. Medical history shall be presented separately for those conditions active/not active at start of study.

Number of patients with medical history and current conditions will be summarized separately using frequency and percentage of patients. Medical conditions with a stop date on or after the dose of first medication are classified as current medical conditions. Medical conditions marked as “Ongoing” are also considered as current medical conditions for presentation. Medical conditions without ongoing ticked or a stop date before the dosing of first study medication are considered to have resolved, and are classified as previous medical history. Medical history details will also be listed. If the stop date is present and before the dosing of first study medication then the event is considered as medical history. In the case of a partially present stop date, if the day is missing then impute the last day of the month unless month is same as month of first dose of study drug then impute last dose date; if the month and day are missing then impute 31st December unless year is the same as first dose date then impute last dose date. Imputed partial stop dates will only be used to assign conditions to medical history or current conditions, and the partially complete date will be used in listings.

The summaries will be presented by primary system organ class (SOC) and preferred term (PT) for each treatment group and in total. A patient will be counted only once within the SOC and PT if the patient has multiple medical conditions within the same SOC and PT.

[redacted]

4.3.3 Prior and concomitant medications

Prior medications are defined as treatments taken and stopped prior to first dose of study treatment. Any medication given at least once between the day of first dose of study treatment and the last day of study visit will be considered as a concomitant medication, including those which were started before study treatment and continued during the study.

Incomplete start date of medications will be imputed (see Appendix 2 for details), this will be done solely for the purpose of assigning a Study Period to the event (see Section 4.1.2) and calculating MTX duration.

If stop date of the medication is clearly prior to the first administration of investigational study medication, then the medication is considered prior, otherwise concomitant. If available data is not sufficient to allow to determine if a medication ended prior to the start of study treatment, it will be considered as NOT ended before the start of study treatment. Missing end dates which are not ongoing will not be imputed, instead these entries will be more conservatively assessed as concomitant. Medications started in Follow-up period (on or after study discontinuation) will not be considered concomitant and analyzed, only listed.

For concomitant medications: if the start date is on or after the first administration in Study Period 2, then medication will be analyzed in Study Period 2 only. If the stop date is clearly before the first administration in Study Period 2, then medication will be analyzed in Study Period 1 only.

Prior and concomitant medications will be coded using WHO Drug Dictionary March 2017 Enhanced, and ATC coding.

4.3.3.1 RA medication

Frequency and percentage of patients received prior RA medication will be summarized by treatment group, Anatomical Therapeutic Chemical (ATC) classification, and PT based on SP1 FAS. Frequency and percentage of patients received concomitant RA medication will be summarized by treatment group, ATC classification, and PT for SP1, SP2, and the entire study separately. Tables will also show the total number and percentage of patients receiving at least one medication of a particular ATC classification.

All RA medication data will be listed.

4.3.3.2 Relevant non-RA medication

Relevant non-RA medication will be summarized and listed in the same way as for RA medication above.

4.4 Treatment regimen, compliance and exposure to drug

4.4.1 Extent of exposure

The extent of exposure to study treatment will be based on the total number of doses administered.

[redacted]

The duration of exposure to study treatment will be summarized by treatment group. In addition, the number of patients with cumulative exposure of at least certain thresholds will be summarized as follows:

- ≥ 2 weeks, ≥4 weeks, ≥ 8 weeks, ≥ 12 weeks, ≥ 18 weeks, and ≥ 24 weeks for SP1 and SP2 separately
- ≥ 2 weeks, ≥4 weeks, ≥ 8 weeks, ≥ 12 weeks, ≥ 18 weeks, ≥ 24 weeks, ≥ 30 weeks, ≥ 36 weeks, ≥ 42 weeks, and ≥ 46 weeks for the entire study

Duration variables (by Study Period and overall) will be derived as given below:

- Duration of exposure to study treatment in SP = date of last dose in SP - date of first dose in TP + 1
- Duration of observation in SP = date of last visit in SP - date of first dose in SP + 1

Date of last visit in SP1 is defined as date of first dose in SP2 if patient is treated in SP2, otherwise as a date of last visit, including End of Study visit.

Duration variables will be presented using summary statistics for continuous variables (as described in Section 2.1.2).

In addition patient exposure years will be presented for each Study Period and overall:

Patient exposure years in a period = (sum of duration of exposure in a period for all patients who received the study drug in a period) / 365.25.

Study drug administration record will be listed as collected in the eCRF.

The analysis of study treatment data will be based on the safety set.

4.4.2 Compliance

Compliance will be calculated using dosages taken before DAS28-CRP Week 24 efficacy assessment for SP1 and before DAS28-CRP Week 48 efficacy assessment for SP2. In case of no Week 24 efficacy assessment all dosages in study period will be taken into account.

Compliance (in percent) to study drug administration in each study period will be calculated as below:

Compliance = 100 × (Number of planned doses in SP – number of missed doses in SP) / Number of planned doses in SP

Twelve doses of administration of study drug administration are planned for a patient in each of Study Period 1 and Study Period 2, twenty-four doses are planned for a patient in the entire study. If the patient discontinues then the expected number of doses will be adjusted using the EOS visit, i.e.

- expected number of dose = int[(date of EOS visit – date of first dose in SP) / 14] + 1
- where int is the integer part of the number, i.e. the number is rounded down to the next integer. For example if the difference of the dates is 13 days then the expected number of doses would be 1, while if the difference in dates is 14 days then the expected number of doses would be 2.

[redacted]

A missing dose of study medication is defined as a totally missing dose (where the date of study drug dose administration is missing), a dose where the full dose was not given.

A summary table will be produced showing the number and percentage of patients who missed any doses during the study period, a number of missed doses and a number of consecutive missed doses, by treatment group and in total.

In addition, compliance will be presented with counts and percentages for categories < 25%, 25 to 50%, 50 to 75% and ≥75% compliance.

The analysis of compliance will be based on FAS.

4.5 Analysis of the primary endpoint

The change from baseline in DAS28-CRP at week 12 will be analyzed using the method described in Section 2.4.1.

The primary analysis will be performed on the Per-protocol set (W12 PPS) which is the most conservative approach to show equivalent efficacy as this analysis set is the "cleanest" sample of the study population to identify differences, if there are any.

The MMRM will include the patient into the analysis as long as there is at least one non-missing DAS28-CRP change from baseline value.

The stratification factors in the primary analysis model will use the data from clinical database as it is most correct data and there are known incorrect stratification assignments in IRT data.

Table 4-1 MMRM Analysis SAS® Code

```
/* Save the LS Means and the differences into datasets */
ods output Diffs = Diffs LSMeans = LSMeans;
proc mixed data=DATASETNAME method=reml covtest;
   class VISIT TREATMENT USUBJID REGION RA_MED BODY_WEIGHT;
   model CHG = BSL VISIT TREATMENT VISIT*TREATMENT REGION
         RA_MED BODY_WEIGHT / solution ddfm=KR;
   repeated VISIT / type=un subject=USUBJID;
   lsmeans TREATMENT / diff CL;        /* Overall treatment effect   */
   lsmeans VISIT*TREATMENT / diff CL;  /* treatment effect per visit */
run;
```

Note: names in capitals are only placeholders for the real dataset or variable names; select the model parameter estimations at visit Week 12 only from the SAS output datasets.

Sensitivity analysis of primary endpoint:

- Repeat the primary analysis based on SP1 FAS
- Repeat the primary analysis using IRT stratification data
- Repeat the primary analysis using the statistical model with treatment as fixed effect and body weight as continuous covariate (i.e., other stratifications like region and prior therapy will be excluded from the model)
- Repeat the primary analysis using the statistical model with treatment as fixed effect only

[redacted]

- Repeat the primary analysis by ADA status up to Week 12 (Positive/Negative)

4.6 Analysis of secondary endpoints

4.6.1 Efficacy evaluation

4.6.1.1 Key secondary endpoint: Time-weighted averaged change from baseline in DAS28-CRP until Week 24

The averaged change from baseline in DAS28-CRP between Week 2 and Week 24 (standardized AUEC approach), will be estimated by a weighted mean across visits. The weights are derived using the standard formula from the trapezoidal rule across the visits. Mean averaged change from baseline from Week 2 to Week 24, standard errors and the two-sided 95%/90% CI for the mean difference between GP2017 and Humira® will be estimated from the ANCOVA model and the 95%/90% CI compared to the pre-specified equivalence margin of [-0.6; 0.6].

The time-weighted averaged change from baseline is defined as the change from baseline of DAS28-CRP in Weeks 2, 4, 12 and 24, weighted based on the time intervals between the actual dates of two consecutive visits in days. The scheduled visits which are outside the allowed time window can be used for the calculation (e.g., the End of Study visit for early discontinued patients). Unscheduled visits or repeat measurements will not be used in the calculation of the time-weighted averaged change from baseline.

In general, the time-weighted averaged change from baseline will be derived using following formula:

$$\text{time weighted averaged CFB} = \frac{\sum_{i=1}^{k} 0.5 \times (c_i + c_{i-1}) \times (d_i - d_{i-1})}{\sum_{i=1}^{k} (d_i - d_{i-1})}$$

Where k = number of timepoints (will be 4 if all DAS28-CRP values up to Week 24 are present)

d_i = study day based on the actual visit dates (see Section 4.1.3)

c_i = change from baseline in DAS28-CRP on d_i, where c_0 refers to baseline and is defined as zero.

If c_m value (where m<k) is missing, c_m shall be set equal to c_{m-1} and $d_i = d_{i-1}$ for the calculation.

The time-weighted averaged change from baseline analysis will be performed using Analysis of Covariance (ANCOVA). The model will include following terms: treatment group, body-weight category, prior systemic therapy, and region as factors and baseline DAS28-CRP value as continuous covariate. In the case that the IRT and eCRF values of the stratification factors differ the correct definition of the stratification factors used for the analyses will be based on the clinical database.

[redacted]

Table 4-2 ANCOVA analysis SAS® Code

```
/* Save the LS Means and the differences into datasets */
ods output Diffs = Diffs LSMeans = LSMeans;
proc mixed data=DATASETNAME covtest;
   class TREATMENT REGION RA_MED BODY_WEIGHT;
   model time-weighted averaged CFB = BSL TREATMENT REGION RA_MED
         BODY_WEIGHT / s ddfm=KR;
   lsmeans TREATMENT / diff CL;
run;
```

Note: names in capitals are only placeholders for the real dataset or variable names

Sensitivity analysis of key secondary endpoint:

- Repeat the key secondary analysis based on SP1 FAS
- Repeat the key secondary analysis using IRT stratification data
- Repeat the key secondary analysis using the statistical model with treatment as fixed effect and body weight as continuous covariate (i.e., other stratifications like region and prior therapy will be excluded from the model)
- Repeat the key secondary analysis using the statistical model with treatment as fixed effect only
- Repeat the key secondary analysis by ADA status up to Week 24 (Positive/Negative)

4.6.1.2 DAS28-CRP and DAS28-ESR scores

The change from baseline in DAS28-CRP and DAS28-ESR at all visits will be analyzed using the same MMRM method as for primary endpoint analysis (see Section 2.4.1).

DAS28-CRP and DAS28-ESR at baseline and weeks 2, 4, 12, 24, 36 and 48, and DAS28-CRP and DAS28-ESR change from baseline at weeks 2, 4, 12, 24, 36 and 48, will be summarized by treatment and visit using descriptive statistics. For DAS28-CRP and DAS28-ESR, the spaghetti plots by treatment and site will be provided, individual patient profile will also be provided.

4.6.1.3 Analysis of EULAR over time

The number and proportion of patients achieving EULAR good and moderate response, as well as EULAR remission criteria in DAS28-ESR at weeks 2, 4, 12, 24, 36 and 48 will summarized and plotted by treatment and visit.

4.6.1.4 Analysis of ACR over time

The number and proportion of patients achieving ACR20/50/70 at weeks 4, 12, 24, 36 and 48 will summarized and plotted by treatment and visit.

The number and proportion of patient achieving EULAR/ACR Boolean remission criteria at weeks 4, 12, 24, 36 and 48, will be summarized and plotted by treatment and visit.

4.6.1.5 Analysis of HAQ-DI

The number and proportion of patients achieved HAQ-DI score with in normal range (≤0.5) at weeks 4, 12, 24, 36 and 48 will be summarized in a table by treatment group and visit.

[redacted]

HAQ-DI score at baseline and weeks 4, 12, 24, 36 and 48 will be summarized of descriptive statistics by treatment group and visit, which will be provided for both the absolute score, change and percent change from baseline, percent change from baseline will also be plotted.

4.6.1.6 Analysis of FACIT

The absolute value, change and percent change from baseline for Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue scale at weeks 4, 12, 24, 36 and 48 will be summarized as a continuous variable by treatment group and visit, percent change from baseline will also be plotted.

4.6.1.7 Analysis of CRP/ESR

The absolute value, change and percent change from baseline for CRP/ESR at weeks 4, 12, 24, 36 and 48 will be summarized by descriptive statistics as a continuous variable by visit and treatment group, percent change from baseline will be plotted.

4.6.1.8 Analysis of rheumatoid factor and anti-CPP antibodies

The number of patients having positive and negative values, and for positive cases a continuous value for rheumatoid factor and anti-CPP at baseline, Week 24 and 48 (for last two also change from baseline) will be summarized by descriptive statistics by treatment group.

4.6.2 Safety evaluation

All safety evaluations will be performed on the Safety set.

4.6.2.1 Adverse Event

AEs will be summarized by presenting the number and percentage of patients having any AE, by SOC and PT.

In addition, the number of events will also be presented for any AE, by primary SOC and for each PT. The average rate of adverse events will also be calculated as the number of events divided by the total number of years of exposure.

Missing and /or incomplete dates for AEs are imputed in a manner resulting in the earliest onset during the study period (see Appendix 1 for details), additionally taking into account that the start date should not be after the stop date. The maximum of the earliest possible start date and the date of first administration of investigational study medication will be imputed for missing dates. If the end date is prior to the date of first administration of investigational study medication then the AE is not considered as a TEAE.

AE summary will be provided for all study periods by treatment groups and for the entire study separately. Following data will be summarized in the overall AE summary presentation.

Total number of TEAEs, patients with at least one TEAE, at least one SAE, at least one treatment related TEAE, at least one severe TEAE, at least one treatment related SAE, patients who discontinued the study due to TEAE, patients requiring study drug interruption and number of patients who died.

[redacted]

4.6.2.1.1 Summary of AEs by severity

Severity is classified as mild/moderate/severe (increasing severity). Treatment emergent AEs starting after the first dose of study medication with a missing severity will be classified as severe. If a patient reports a TEAE more than once within that SOC/PT, the TEAE with the highest severity will be presented.

4.6.2.1.2 Summary of study treatment related AEs.

Relationship, as indicated by the investigator, is classed as suspected/not suspected/not applicable (before study drug administration). A related TEAE is defined as a TEAE with a relationship to study medication as "suspected" to study medication. Treatment emergent AEs with a missing relationship to study medication will be regarded as "suspected" to study medication. If a patient reports the same TEAE more than once within that SOC/PT, the TEAE with the worst case relationship to study medication will be presented.

4.6.2.1.3 Summary of AEs of special interest

Selected AEs, considered under the heading of potential and identified risks will be presented under summary of AEs of special interest. Definition of an AE as an AE of special interest is given in Appendix 3.

In addition separate summaries will be provided for death, serious AEs (SAEs), other significant AEs leading to discontinuation by treatment group and overall.

These summaries will be presented separately for all study periods by treatment group and overall.

4.6.2.1.4 Injection site reactions

All AEs marked as ISRs in the eCRF will be summarized by presenting the number and percentage of patients having any ISR and by signs/symptoms. Injection site reactions will also be presented by severity. In addition, the number of events for ISRs will also be presented overall for any ISR and by sign/symptom. If a patient reported more than one identical ISR, the ISR with the greatest severity will be presented. If a patient reported more than one identical ISR, the patient will be counted only once with the greatest severity. If the severity is missing for an ISR, severity will be classified as severe. ISRs will be presented separately for all study periods by treatment groups and overall. In addition, ISRs which are considered SAEs will be presented as counts and percentages separately.

4.6.2.2 Laboratory data

Results from the central laboratory will be included in the reporting of this study. The summaries presenting a number of patients with new notable abnormalities (according to the protocol), as well as shift tables by toxicity grades will be presented for hematology and serum chemistry by treatment group for each Study Period. Standard units will be used for presentation. Box plots of parameters of interest will be done by treatment and study period. The laboratory results below the LLOQ will be imputed as zero, the laboratory results above the upper detection limit will be imputed as upper detection limit.

[redacted]

All laboratory data (including urinalysis) will be listed. Notable abnormalities will be listed in addition.

4.6.2.3 Vital Signs

Vital signs measurement includes blood pressure (systolic and diastolic), pulse and temperature. The summary presenting a number of patients with new notable abnormalities (according to normal range criteria) will be presented by treatment group for each study period. All information collected will be listed. Notable abnormalities will be listed in addition.

4.6.2.4 ECG

All ECG results will be shown in a listing.

4.6.2.5 Physical examination

Physical examination results will be shown in listing.

4.6.3 Immunogenicity

Summary statistics (frequency and percentage) by treatment group and overall and sampling day will be provided for the percent of patients with ADA formation at baseline, weeks 2, 4, 12, 24, 36 and 48 or End of Study visit.

ADA positive patients at baseline will not be considered for the analysis of immunogenicity (development of ADA) at post-baseline assessments. A patient is considered as ADA positive during the study if the patient had at least one post baseline ADA positive sample otherwise the patient is regarded as ADA negative. Note this will also include any unscheduled visits during the Study Period.

A transient ADA response occurs where a post baseline positive ADA response occurs but is followed by a negative ADA response during the study.

The titer and the concentration of confirmed positive results will be reported. In addition, confirmed positive ADA samples will be analyzed for their neutralization potential in a neutralizing antibody (NAb) assay.

The incidence of ADA positive patients, patients with a transient response and patients with neutralizing ADAs will be summarized by treatment arms.

All ADA data will be listed (including titer and concentration).

The SAF will be used to summarize the data.

4.6.4 Pharmacokinetic evaluations

Through adalimumab serum concentrations at Baseline, Weeks 2, 4, 12, 24, 30, 36 and 48 or End of Study visit will be summarized with descriptive statistic by treatment group and overall. Concentrations will be listed together with ADA results.

4.7 Handing of missing data in the analysis

Refer to Section 2.3.1.1, Appendix 1, and Appendix 2.

[redacted]

5 Clinical Study Report – Appendix 16.1.9 Documentation of statistical methods

Appendix 16.1.9 of the CSR will contain the last SAP versions before database lock and the full SAS® outputs of the MMRM analyses of DAS28-CRP and DAS28-ESR as well as the ANCOVA analyses of the time-weighted averaged change from baseline approaches.

6 Appendices

6.1 Appendix 1: Adverse event partial date imputation

6.1.1 Adverse event end dates

- If the day is missing, then impute the last day of the month
- If the day and month are both missing, then impute the date with the day and month of the last available assessment date for the patient.
- If the end date is completely missing, the event will be assumed to be 'Ongoing'

6.1.2 Adverse event start date

The algorithm only applies to event that has partial start date. Please note that **completely missing start dates** will not be imputed.

Note: Imputed dates will be used in the derivation of treatment emergent AEs and assignment of study periods. The actual partial dates will be reported in the listings.

The following matrix explains the logic behind the imputation.

	MON MISSING	**MON < TRTM**	**MON = TRTM**	**MON > TRTM**
YYYY < TRTY	**(D)** Before Treatment Start	**(C)** Before Treatment Start	**(C)** Before Treatment Start	**(C)** Before Treatment Start
YYYY = TRTY	**(B)** Uncertain	**(C)** Before Treatment Start	**(A)** Uncertain	**(A)** After Treatment Start
YYYY > TRTY	**(E)** After Treatment Start	**(A)** After Treatment Start	**(A)** After Treatment Start	**(A)** After Treatment Start

The following table is the legend to the logic matrix.

	Day	**Month**	**Year**
Partial Adverse Event Start Date	Not used	MON	YYYY
Treatment Start Date (TRTSTD)	Not used	TRTM	TRTY
Treatment in SP2 Start Date (TRT2STD)	Not used	TRT2M	TRT2Y

Relationship	
Before Treatment Start	Partial date indicates AE start date **prior** to Treatment Start Date
After Treatment Start	Partial date indicates AE start date **after** Treatment Start Date
Uncertain	Partial date **insufficient to determine** relationship of AE start date to Treatment Start Date
Imputation Calculation	
NC / Blank	No convention
(A)	MAX (01MONYYYY, TRTSTD)
(B)	TRTSTD
(C)	15MONYYYY
(D)	01JULYYYY
(E)	01JANYYYY

In cases in the categories (A) and (E) when it is unclear due to the partial date if the AE started in SP1 or SP2 (YYYY = TRT2Y and MON=TRT2M in SP2) it will be considered to have started in SP2 and AE start date will be imputed as TRT2STD.

[redacted]

6.2 Appendix 2: Concomitant medication partial date imputation

6.2.1 Concomitant medication end date

- If the day is missing, then impute the last day of the month
- If the day and month are both missing, then impute the date with the day and month of the last available assessment date for the patient.
- If the end date is completely missing, the event will be assumed to be 'Ongoing'

6.2.2 Concomitant medication start date

This algorithm only applies to event that has partial start date. Please note that **completely missing start dates** will not be imputed.

Note: Imputed dates will be used in the derivation of prior medication or concomitant medication, and assignment of study periods. The actual partial dates will be reported in the listings. If the imputed concomitant medication (CMD) start date is after CMD end date, it will be set to CMD end date.

The following matrix explains the logic behind the imputation.

	MON MISSING	**MON < TRTM**	**MON = TRTM**	**MON > TRTM**
YYYY < TRTY	**(D)** Before Treatment Start	**(A)** Before Treatment Start	**(A)** Before Treatment Start	**(A)** Before Treatment Start
YYYY = TRTY	**(C)** Uncertain	**(A)** Before Treatment Start	**(C)** Uncertain	**(B)** After Treatment Start
YYYY > TRTY	**(E)** After Treatment Start	**(B)** After Treatment Start	**(B)** After Treatment Start	**(B)** After Treatment Start

The following table is the legend to the logic matrix.

	Day	**Month**	**Year**
Partial CMD Start Date	Not used	MON	YYYY
Treatment Start Date (TRTSTD)	Not used	TRTM	TRTY
Relationship			
Before Treatment Start	Partial date indicates CMD start date **prior** to Treatment Start Date		
After Treatment Start	Partial date indicates CMD start date **after** Treatment Start Date		
Uncertain	Partial date **insufficient to determine** relationship of CMD start date to Treatment Start Date		
Imputation Calculation			
NC / Blank	No convention		
(A)	15MONYYYY		
(B)	MAX (01MONYYYY, TRTSTD)		
(C)	TRTSTD		
(D)	01JULYYYY		
(E)	01JANYYYY		

According to this imputation, partial start dates during the treatment will be imputed with the earliest possible date, thus if from the partial date it is unclear if the medication started in SP1 or SP2 then it will be considered to have started in SP1. Note that when a partial start date clearly belongs to SP2 then the medication will be assigned to SP2 only.

[redacted]

6.3 Appendix 3: Adverse events of special interest

- Infections
  - Tuberculous infections (HLT)
  - Atypical mycobacterial infections (HLT)
  - Pneumonia (PT)
  - Hepatitis B (PT)
  - Acute Hepatitis B (PT)
  - Hepatic viral infections (HLT)
  - Hepatitis C (PT)
  - Acute Hepatitis C (PT)
  - Sepsis, bacteraemia, viraemia and fungaemia NEC (HLT)
  - Listeriosis (PT)
  - Histoplasmosis (PT)
  - Legionella infection (PT)
  - Pneumonia legionella (PT)
  - Fungal infectious disorders (HLGT)
    - Pneumocystis infections (HLT)
    - Aspergillus infections (HLT)
    - Herpes viral infections (HLT)
    - Blastomycosis (HLT)
    - Coccidioidimycosis (HLT)
  - Arthritis bacterial (PT)
  - Erysipelas (PT)
  - Cellulitis (PT)
  - Diverticulitis (PT)
  - Pyelonephritis (PT)
- Malignancies
  - Neoplasms benign, malignant and unspecified (incl cysts and polyps) (SOC)
- Allergic reactions
  - Angioedema and urticaria (HLGT)
  - Urticaria (PT)
  - Hypersensitivity (PT)
  - Anaphylactic reaction (PT)
  - Drug hypersensitivity (PT)
  - Allergic Bronchospasm (PT)
- Immune system disorders/Autoimmune events
  - Acute and chronic sarcoidosis (HLT)
  - Vasculitides (HLT)

[redacted]

  - Vasculitides NEC (HLT)
  - Autoimmune pancytopenia (PT)
  - Autoimmune hepatitis (PT)
  - Autoantibody positive (PT)
  - Lupus-like syndrome (PT)
  - Interstitial lung disease (PT)
- Neurological events
  - Demyelinating disorders
    - Multiple sclerosis (PT)
    - Optic neuritis (PT)
    - Guillain-Barre syndrome (PT)
- Hematological reactions
  - Pancytopenia (PT)
  - Thrombocytopenia (PT)
  - Anaemia (PT)
  - Aplastic Anaemia (PT)
  - Leukopenia (PT)
  - Neutropenia (PT)
  - White blood cell count decreased (PT)
  - Agranulocytosis (PT)
  - Leukocytosis (PT)
  - Idiopathic thrombocytopaenic purpura (PT)
- Congestive Heart Failure
  - Cardiac failure congestive (PT)
  - Hypertension (PT)
  - Arrhythmia (PT)
- Gastrointestinal disorders
  - Gastrointestinal haemorrhage (PT)
  - Pancreatitis (PT)
  - Intestinal perforation (PT)
- Hepato-biliary disorders
  - Hepatic enzyme increased (PT)
  - Cholecystitis (PT)
  - Cholelithiasis (PT)
  - Hepatic steatosis (PT)
  - Hepatic failure (PT)
- Renal and urinary disorders
  - Renal impairment (PT)

[redacted]

- Respiratory, thoracic and mediastinal disorders
  - Asthma (PT)
  - Interstitial lung disease (PT)
  - Pulmonary embolism (PT)
  - Pleural effusion (PT)
  - Pulmonary fibrosis (PT)

6.4 Appendix 4: Medications of special interest

The following medications in Table 6-1 below are used to identify the different classes of DMARD medication.

Table 6-1 ATC codes and preferred terms for DMARDs

Class	ATC third/fourth level	ATC Code	Preferred term
Non-biologic (other) DMARDs	Nitrogen mustard analogues	L01AA	Cyclophosphamide
	Protein Kinase Inhibitors	L01XE	Tyrosin kinase inhibitors, Fostamatinib
	Selective immunosuppressants	L04AA	Leflunomide, Tofacitinib, Mycophenolic acid, Teriflunomide, Apremilast
	Calcineurin inhibitors	L04AD	Ciclosporine
	Other immunosuppressants	L04AX	Azathioprine
	Other anti-inflammatory and antirheumatic agents, non-steroids	M01AX	Hydroxychloroquine, Hydroxychloroquine sulfate, Sulfasalazine
	Gold preparations	M01CB	Aurothioglucose, Gold, Sodium aurothiomalate, Auranofin
	Aminoquinolines	P01BA	Chloroquine or Chloroquine phosphate, Hydroxychloroquine
	Penicillamine and similar agents	M01CC	Penicillamine, Bucillamine
	Investigational Drug	V98	• investigational product or placebo (dissociated agonis of glucocorticoid receptor) (PF-04171327)[1] • investigational product or placebo(S1P-lyase inhibitor) (LX3305)[1] • investigational product or placebo (JAK inhibitor)(ASP-015K)[1]
Methotrexate	Other immunosuppressants or Other specific antirheumatic agents	L04AX or M01CX	Methotrexate or Methotrexate sodium

[redacted]

Class	ATC third/fourth level	ATC Code	Preferred term
Biologic DMARDs	Interleukins	L03AC	Interleukins
	Selective immunosuppressants	L04AA	Abatacept, Efalizumab, Alefacept, Belimumab,
	Tumor Necrosis Factor Alpha (TNF-) Inhibitors	L04AB	Etanercept, Golimumab, Infliximab
	Interleukin inhibitors	L04AC	Tocilizumab, Anakinra
	Monoclonal antibodies	L01XC	Rituximab, Monoclonal Antibodies
	Other specific antirheumatic agents	M01CX	Tregalizumab
	Investigational Drug	V98	• investigational product (biologic, IL-17 receptor inhibitor) (AMG827) • AMG 827 (anti - IL17) • antagonist IL-17 (CNTO 6785) • clinical trial nn38226-3613 Phase 2b (IL-20)

[1]: treatment in clinical trials (active substance or placebo) was assessed as active substance

References

Burmester G-R, Kivitz AJ, Kupper H, et al (2015) Efficacy and safety of ascending methotrexate dose in combination with adalimumab: the randomised CONCERTO trial. Ann Rheumatic Dis; 74:1037-44

Ibrahim F, Tom BDM, Scott DL, et al (2016) A Systematic review of randomized controlled trials in rheumatoid arthritis: the reporting and handling of missing data in composite outcomes. Trials; 17:272

Wells G, Becker J-C, Teng J, et al (2009) Validation of the Disease Activity Score 28 (DAS28) and EULAR response criteria based on CRP against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on ESR. Ann Rheum Dis; 68:954-60

[redacted]