Non-Interventional Study Protocol
A3921342

BASELINE VARIABLES PREDICTING TREATMENT RESPONSE AT 6 MONTHS IN ADULT RHEUMATOID ARTHRITIS PATIENTS TREATED WITH TOFACITINIB 5MG BID IN A NON-INTERVENTIONAL SETTING

Statistical Analysis Plan
(SAP)

Version: 1.0

Author: PPD

Date: 28-Jan-2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Definition
ACPA	Anti-citrullinated protein antibodies
AE	Adverse Event
CDAI	Clinical Disease Activity Index
CRP	C-Reactive Protein
csDMARD	Conventional synthetic Disease Modifying Antirheumatic Drugs
DAS	Disease Activity Score
DMARD	Disease Modifying Antirheumatic Drugs
EQ-5D	EuroQol five dimensions questionnaire
ESR	Erythrocyte sedimentation rate
FACIT	Functional Assessment of Chronic Illness Therapy
HAQ-DI	Health Assessment Questionnaire – Disability Index
JAK	Janus kinase (JAK)
LDL	Low-density lipoprotein,
MTX	Methotrexate
MMRM	Mixed model repeated measures
NI	Non-Interventional
NRI	Non-responder Imputation
PhGA	Physician's Global Assessment
PRO	Patient reported outcome

PtGA	Patient's global assessment of health
RA	Rheumatoid arthritis
RF	Rheumatoid Factor

Abbreviation	**Definition**
SAE	Serious Adverse Events
SAP	Statistical Analysis Plan
SAS	Statistical Analysis Software
SD	Standard deviation
SDAI	Simplified Disease Activity Index
SJC	Swollen joint count
TJC	Tender joint count
VAS	Visual Analogue Scale
WPAI	Work Productivity and Activity Impairment

1. AMENDMENTS FROM PREVIOUS VERSION(S)

This is the first version of the statistical analysis plan (SAP). There are no changes to the analyses specified in the protocol.

2. INTRODUCTION

This SAP is based on protocol A3921342 Protocol Version 2 dated 14th January 2020

Note: in this document any text taken directly from the non-interventional (NI) study protocol is ***italicised***.

Rheumatoid arthritis (RA) is a chronic, autoimmune disease characterized by joint inflammation and destruction, leading to progressive disability and adverse psychological effects with severe impact on patients' quality of life.

The chronic nature of the disease deems necessary the long-term use of medication. Thus, RA represents significant health and socioeconomic burdens for the individual subject, healthcare systems and society, with indirect costs (productivity loss, absenteeism/disability, early retirement) accounting for up to ¾ of total costs for working patients.

As there is currently no cure for RA, the target of treatment strategies is to control disease activity, alleviate signs and symptoms, maintain physical function, optimize quality of life, reduce the rate of joint damage, and, if possible, induce and maintain complete remission. Although there are several effective treatments available nowadays, mainly biologic DMARDs, many patients fail to achieve or maintain the desired response and are treated sequentially with different drugs leading to increased costs, risk of toxicity and suboptimal effectiveness. Therefore, the ambition for personalized treatment has emerged and investigation to identify predictive tools of treatment response is a highly active area of research.

Tofacitinib (CP-690,550) is an oral Janus kinase (JAK) inhibitor that targets inflammation by reducing pro-inflammatory cytokine signaling and production.

The efficacy and safety evaluation of Tofacitinib for the treatment of RA is based on a comprehensive clinical development program including six randomized, double blind, multicenter phase III studies in adult subjects with active in rheumatoid arthritis (RA). In these trials Tofacitinib consistently reduces signs and symptoms of RA, improves physical function and other subject-reported outcomes such as fatigue, pain and health-related aspects of quality of life in subjects with moderate to severe RA. Combined with its inhibition of the progression of structural damage, the development program has demonstrated Tofacitinib as an effective targeted synthetic disease-modifying antirheumatic drug (DMARD) in treating RA.[i] *Long-term extension and Phase 3b/4 post authorization safety studies are aimed to demonstrate sustained efficacy and a consistent safety profile as seen in the Phase 2 and 3 controlled clinical trials.*

However, these trials may have limitations reflecting real world situations. It has been estimated that only 21-33% of patients documented in registries, reflecting more accurately routine care, would have met eligibility criteria of major trials as they often exclude patients with co-morbidities. Also treatment of patients in daily clinical practice may vary from strict regulations in clinical trials. Therefore additional long-term data in registries and non-interventional studies of a European population would be helpful to further assess treatment outcomes in a real-world setting.

Furthermore, studies exploring the association between clinical and biochemical parameters with the efficacy of Tofacitinib treatment are scarce. Prior, analyses based on P123 and LTE Tofacitinib data have shown: a) early achievement of clinical disease control seems to determine longer term treatment effect with Tofacitinib in MTX-naïve patients; b) Tofacitinib treatment response at 6m may be higher for patients with higher disease activity at baseline based on CRP levels particularly for those with inadequate response to biologic DMARDs.. These results point out the need for more research and further investigation in order to help identify biomarkers that can predict response to Tofacitinib treatment in patients with RA.

The impact of RA on work can be profound since permanent work disability (inability to continue working), is common among patients with RA. In addition to the consequences for the patient, such as a decreased quality of life, work disability also leads to high costs. It is estimated that approximately one-third of the total cost for patients with RA is caused by production losses. This implies that at-work productivity loss is an important concern, since work hours are not only lost incidentally through sick leave, but also more structurally and profoundly by at-work productivity loss. As so far data on the impact of Tofacitinib on work productivity are still limited, this study will assess the effect Tofacitinib on work productivity and non-work activities.

2.1. Study Design

This is a 24-month, prospective, non-interventional, multi-center study with primary aim to identify clinical and/or biochemical factors that predict achievement of remission or Low Disease Actvity (LDA) *in patients with RA treated with Tofacitinib in Greece. Eligible subjects will be followed from the date of first Tofacitinib prescription until study end, death, or discontinuation, or loss to follow-up (whichever comes first) for the occurrence of the endpoints of interest.*

In this study, the medicinal product(s) are prescribed in the usual manner in accordance with the terms of the marketing authorisation.

The assignment of the patient to a particular therapeutic strategy is not decided in advance by this protocol, but falls within current practice and the prescription of the medicine should be clearly separated from the decision to include the patient in the study.

No additional diagnostic or monitoring procedures will be applied to the patients and follow-up visits are captured as part of normal medical practice. Epidemiological methods will be used for the analysis of collected data.

Figure 1. Study Design

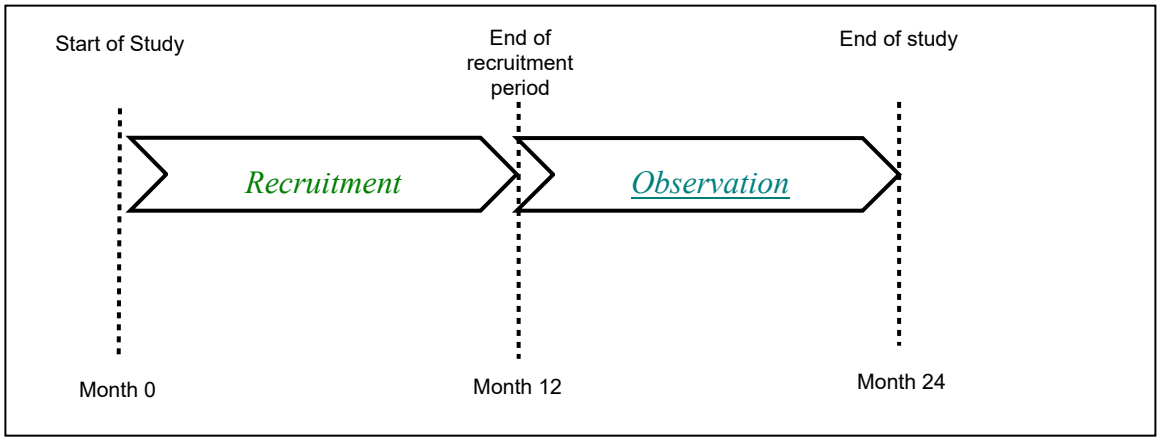

Table 1 shows the key assessments and their associated time-points for data collection.

Table 1. Schedule of Activities

Study Week	**Baseline** (Enrolment)	**Month 3**	**Month 6**	**Month 12**
Visit Number	1	2	3	4
Informed Consent	x			
Anamnesis / Medical History	x			
Demographic data	x			
Educational level	x			
Working status	x			x
Inclusion/exclusion criteria	x			
Vaccination status	x			
Treatment/Medication				
Prior RA drug treatment	x			
Concomitant RA drug treatment	x	x	x	x
Co-medication	x	x	x	x
Safety				
Inflammatory markers	x	x	x	x
Documentation of Comorbidities	x			
Documentation of AE		x	x	x
Efficacy				

Number of swollen and tender joints	x	x	x	x
Morning stiffness	x	x	x	x
Patients Global Assessment of Arthritis	x	x	x	x
Physicians Global Assessment of Arthritis	x	x	x	x
Patient Assessment of Arthritis Pain	x	x	x	x
Health Assessment Questionnaire – Disability Index (HAQ-DI)	x	x	x	x
EuroQol EQ-5D Health State Profile	x	x	x	x
Work Productivity and Activity Impairment (WPAI) Questionnaire	x	x	x	x
FACIT (Functional Assessment of Chronic Illness Therapy) – Fatigue	x	x	x	x
Overall satisfaction with treatment	x			x
Medical visits		x	x	x
Diagnostic tests		x	x	x
Hospitalizations		x	x	x

2.2. Study Population

The study population consists of patients with RA treated with Tofacitinib in Greece. *Overall, about 200 patients in 19 centres are to be included in this non-interventional study*.

The planned recruitment period is 12 months. With a planned observation duration of 12 months per patient, the entire study would thus last for 24 months. A patient will be followed up until the end of the planned study duration, regardless of any missing intermediate visits, as well as treatment changes during the course of follow-up.

2.3. Treatment/cohort labels

The treatment labels for analyses are as follows:

Treatment	Treatment Label
Tofacitinib (as monotherapy and combination)	Tofacitinib

2.4. Study Objectives

The primary objective of this study is to identify clinical and/or biochemical factors that predict achievement of remission or LDA in patients with RA treated with Tofacitinib. Logistic regression will be used to investigate the impact of potential predictive factors on remission or LDA at month 6 of treatment.

The secondary objectives of the study are:

- *To describe the treatment patterns of RA patients prescribed Tofacitinib in a realworld setting;*
- *Assess the effect of treatment on patient quality of life and physical function;*
- *Estimate resource utilization and costs in patients with rheumatoid arthritis (RA) treated with Tofacitinib in Greece.*

The secondary objectives will be assessed using descriptive methods unless otherwise stated. Quality of life, measured using the EQ-5D, Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale and the Work Productivity and Activity Impairment (WPAI) questionnaire, will be analysed using mixed model repeated measures (MMRM) models.

3. INTERIM ANALYSES

An interim analysis will not be performed in this study.

4. HYPOTHESES AND DECISION RULES

This study is CCI with no treatment comparisons. There are no formal hypotheses or decision rules.

5. ANALYSIS SETS/POPULATIONS

5.1. Full Analysis Set

The Full Analysis Set is defined as patients who receive at least one dose of Tofacitinib and have at least one set of post-baseline measurements. This population will be used for all efficacy analysis.

5.2. Safety Analysis Set

The Safety Analysis Set is defined as all patients who receive at least one dose of Tofacitinib. This population will be used for all safety analysis.

5.3. Subgroups

The following subgroups will be considered for health resources and costs summaries and analyses. Cross-classification tables may also be presented combining the following subgroups depending on data availability.

- **Education Status**
  - Primary;
  - Secondary; Tertiary (higher); Other.

- **Working Status**
  - Student;
  - Part-time employed;
  - Full time employed;
  - Self-employed;
  - Not working but looking for work;
  - Not working not looking for work; Not working due to disability/illness; Retired.

- **Disease Severity**
  - **As defined by DAS28-4(CRP):**
  - DAS28 >5.1 corresponds to a high disease activity;
  - 3.2<=DAS28<=5.1 corresponds to a moderate disease activity; 2.6<=DAS28<3.2 corresponds to a low disease activity; DAS28 value <2.6 corresponds to remission.

- **Treatment**
  - Tofacitinib as monotherapy;
  - Tofacitinib in Combination with another DMARD; Other DMARD.
    (See Section 11 A1.2 for definitions of DMARD).

- **Treatment Pattern**
  - Tofa mono;
  - Tofa combo;
  - Tofa mono to combo; Tofa combo to mono;
  - Tofa to other treatment.

- **Co-morbidities**
- The three most common co-morbidities will be considered.

- **HAQ-DI at Baseline**
  - HAQ-DI at Baseline will be categorised as follows:
  - <0.5;
  - [0.5-1.0);
  - [1.0-1.5);
  - [1.5-2);
  - [2-2.5); ≥2.5.

- **EQ-5D at Baseline**
  - The following categories will be used:
  - <0.4;
  - (0.4-0.6];
  - (0.6-0.8]; (0.8-1.0].

6. ENDPOINTS AND COVARIATES

6.1. Efficacy/Effectiveness Endpoint(s)

Baseline measurements for all endpoints are the last measurement before Tofacitinib treatment, up to 28 days before starting Tofacitinib. Change from baseline at a timepoint will be calculated as:

Measurement at visit – Measurement at baseline

6.1.1. Primary Endpoints

The primary endpoints are the rate of remission and low disease actvitiy (LDA) at month 6 of treatment. These are assessed using the 28-joint Disease Activity Score C-reactive protein (DAS28-4 (CRP)), which is a composite endpoint, calculated using 4 variables (represented by '-4' in the name).

Remission and LDA are defined as DAS28-4 (CRP) <2.6 and DAS28-4 (CRP) <3.2, respectively.

The DAS28-4 (CRP) will be calculated as follows:

DAS28-4 (CRP) =0.56 * √ (TJC) + 0.28 * √ (SJC) + 0.36 * ln(CRP+1) + 0.014*PtGA + 0.96

Where CRP is in mg/l, TJC = tender joint count, SJC=swollen joint count, ln = nautral logarithm and PtGA is the Patient's Global Assessment of Health in mm.

6.1.2. Secondary Endpoints

6.1.2.1. Remission or LDA at Month 12

The rate of remission and LDA at month 12 of treatment, assessed by DAS28-4 (CRP) < 2.6 and DAS28-4 (CRP) <3.2, respectively. DAS28-4 (CRP) will be calculated as for the primary endpoint.

6.1.2.2. Change from baseline in DAS28-4 (ESR) and DAS28-4 (CRP)

The 28-joint Disease Activity Score erthrocyte sedimentation rate (DAS28-4 (ESR)) is a composite endpoint, calculated using 4 variables (represented by '-4' in the name).

DAS28-4 (ESR) will be derived as follows:

DAS28-4 (ESR) = 0.56*√(TJC28) + 0.28*√(SJC28) + 0.70*In(ESR) + 0.014*PtGA

Where TJC = tender joint count, SJC=swollen joint count, ln = natural logarithm, ESR is measure in mm/ hour and PtGA is the Patient's Global Assessment of Health in mm.

DAS28-4 (CRP) is defined in Section 6.1.1.

Change from Baseline in DAS28-4 (ESR) and DAS28-4 (CRP) will be calculated for each timepoint.

6.1.2.3. Change from Baseline in HAQ-DI

The Health Assessment Questionnaire – Disability Index (HAQ-DI) assesses the degree of difficulty a patient has experienced during the past week in 8 domains/categories of daily living activities: dressing and grooming, arising, eating, walking, hygiene, reach, grip, and other activities. Each activity category consists of 2-3 items. For each question in the questionnaire, the level of difficulty is scored from 0 to 3 with 0 representing "no difficulty," 1 as "some difficulty," 2 as "much difficulty," and 3 as "unable to do".

The HAQ-DI will then be derived using the Scaling and Scoring Guidelines v4.0 (July 2011). This can be found in Section 11.2.

The HAQ-DI is calculated as follows:

- For each sub-category take the maximum score of all questions in that subcategory (score of 0-3).

- If aids were used for the sub-category a score of 0 or 1 is increased to 2 to more accurately represent underlying disability (scores of 3 are not modified).
- The average of the 8 categories is then calculated to obtain an overall HAQ-DI score of 0-3.

Section 7.1.1 provides the rules for deriving the HAQ-DI if there is missing data.

Change from Baseline in the HAQ-DI will be calculated for each time-point.

HAQ-DI will also be categorized using the subgroups outlined in Section 5.3.

6.1.2.4. Rate of Remission

Patient remission will be derived in four separate ways:

1. **SDAI Remission**

   The Simplified Disease Ativity Index (SDAI) will be derived at each time-point using the formula:

   SDAI=(TJC28) + (SJC28) + PhyGA + PtGA + CRP

   Where TJC = tender joint count, SJC = swollen joint count, PhyGA is the Physician's Global Assessment in cm, PtGA is the Patient's Global Assessment of Health in cm and CRP is C-reactive Protein in mg/dL.

   SDAI Remission will then be derived as follows:

   SDAI_Remission = 1, if SDAI $\leq$3.3
   SDAI_Remission = 0, otherwise.

   The rate of remission based on SDAI will then be derived at month 3, 6 and 12.

2. **CDAI Remission**

   The Clincal Disease Ativity Index (CDAI) will be derived at each time-point using the formula:

   CDAI=(TJC28) + (SJC28) + PhyGA + PtGA

   Where TJC = tender joint count, SJC=swollen joint count, PhyGA is the Physician's Global Assessment in cm and PtGA is the Patient's Global Assessment of Health in cm.

   CDAI Remission will then be derived as follows:

CDAI_Remission = 1, if CDAI ≤ 2.8
CDAI_Remission = 0, otherwise.

The rate of remission based on CDAI will then be derived at month 3, 6 and 12.

3. **DAS 28-4 (ESR) Remission**

DAS 28-4 (ESR) will be derived at each time-point as outlined in Section 6.1.2.2.

DAS 28-4 (ESR) Remission will then be derived as:

DAS 28-4 (ESR)_Remission = 1, if DAS28-4 (ESR) <2.6
DAS 28-4 (ESR)_Remission = 0, otherwise.

The rate of remission based on DAS 28-4 (ESR) will then be derived at month 3, 6 and 12.

4. **DAS28-4 (CRP) Remission**

DAS 28-4 (CRP) will be derived at each time-point using the formula in Section 6.1.1.

DAS 28-4 (CRP) Remission will then be derived as:

DAS 28-4 (CRP)_Remission = 1, if DAS28-4 (CRP) < 2.6
DAS 28-4 (CRP)_Remission = 0, otherwise.

The rate of remission based on DAS 28-4 (CRP) will then be derived at month 3, 6 and 12.

6.1.2.5. Rate of LDA

LDA for each patient will be derived in four separate ways:

1. **SDAI LDA**

SDAI will be derived at each time-point using the formula in Section 6.1.2.4.

A patient is considered as having LDA if SDAI≤11.

The rate of LDA based on SDAI wil then be derived at month 3, 6, 12.

2. **CDAI LDA**

CDAI will be derived at each time-point using the formula in Section 6.1.2.4.

A patient is considered as having LDA if CDAI≤10.

The rate of LDA based on CDAI will then be derived at month 3, 6, 12.

3. **DAS 28-4 (ESR) LDA**

   DAS 28-4 (ESR) will be derived at each time-point using the formula in section 6.1.2.2.

   A patient is considered as having LDA if DAS 28-4 (ESR) <3.2

   The rate of LDA based on DAS 28-4 (ESR) will then be derived at month 3, 6, 12.

4. **DAS 28-4 (CRP) LDA**

   DAS 28-4 (CRP) will be derived at each time-point using the formula in section 6.1.1.

   A patient is considered as having LDA if DAS 28-4 (CRP) <3.2

   The rate of LDA based on DAS 28-4 (CRP) will then be derived at month 3, 6, 12.

6.1.2.6. HAQ-DI Response

A HAQ-DI response is defined as a decrease from baseline of at least 0.22 .

The percentage of patients achieving a HAQ-DI response at each time-point (month 3, 6 and 12) will be derived.

6.1.2.7. Change from Baseline in EQ-5D

The EQ-5D is a standardised instrument used to measure quality of life. It is based on five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has three responses and the patient is asked to select the response that best describes them. The responses are scored 1-3 as shown in Table 2:

Table 2. EQ-5D

Dimension	Response	Score
Mobility	I have no problems in walking about	1
	I have some problems in walking about	2
	I am confined to bed	3
Self-care	I have no problems with self-care	1

	I have some problems washing or dressing myself	2
	I am unable to wash or dress myself	3
Usual activities (eg, work, study, housework, family or leisure activities)	I have no problems with performing my usual activities	1
	I have some problems with performing my usual activities	2
	I am unable to perform my usual activities	3
Pain/discomfort	I have no pain or discomfort	1
	I have moderate pain or discomfort	2
	I have extreme pain or discomfort	3
Anxiety/depression	I am not anxious or depressed	1
	I am moderately anxious or depressed	2
	I am extremely anxious or depressed	3

The score for each dimension is weighted in accordance with Table 3. These weightings are based on UK data (Dolan et al 1995).

Table 3. EQ-5D weightings

EQ-5D Dimension	Score = 1	Score = 2	Score = 3
Mobility	0	0.069	0.314
Self-Care	0	0.104	0.214
Usual Activities	0	0.036	0.094
Pain/Discomfort	0	0.123	0.386
Anxiety/Depression	0	0.071	0.236

The following algorithm is then applied to calculate the EQ-5D Total Score:

1. If **all** five EQ-5D dimensions have a score of 1 then the EQ-5D Total Score is 1.

2. If **any** of the five EQ-5D dimensions have a score of 3, then the EQ-5D Total Score is:

$$1 - 0.081 - \left(\sum_{1}^{5} weighted\ dimension\ score\right) - 0.269$$

3. If **none** of the five EQ-5D dimensions has a score of 3, then the EQ-5D Total Score is:

$$1 - 0.081 - \left(\sum_{1}^{5} weighted\ dimension\ score\right)$$

Missing weighted dimension scores are imputed as outlined in Section 7.1.2.

The EQ-5D score will be derived for all time points and the change from baseline calculated.

The EQ-5D score at each visit will also be categorised into the following subgroups:

- <0.4;
- (0.4-0.6];
- (0.6-0.8];
- (0.8-1.0].

6.1.2.8. Change from Baseline in FACIT-fatigue

The FACIT-Fatigue score (V4.0 16Nov2007) is derived by taking the sum of the scores for the 13 questions in the instrument, resulting in a score between 0 and 52. The questions consider the patients functionality over the 7 days prior to completing the questionnaire. Higher scores represent better patient status (less fatigue).

For some of the questions, the scale indicates a higher score is better and for others, a lower score is better. Therefore, in order to ensure that higher scores represent better status, the items marked with a '–' in table 4 below will be reversed by subtracting the response from "4" before summing all subscale items to derive the FACIT-Fatigue score. **Table 4. FACIT-Fatigue**

Item No.	Question	+/-
HI7	I feel fatigued	-
HI12	I feel weak all over	-
An1	I feel listless ("washed-out")	-
An2	I feel tired	-

An3	I have trouble starting things because I am tired	-
An4	I have trouble finishing things because I am tired	-
An5	I have energy	+
An7	I am able to do my usual activities	+
An8	I need to sleep during the day	-
An12	I am too tired to eat	-
An14	I need help doing my usual activities	-
An15	I am frustrated by being too tired to do the things I want to do	-
An 16	I have to limit my social activity because I am tired	-

See Section 7.1.3 for details on the rules for deriving the FACIT-Fatigue score if there is missing data.

The FACIT-Fatigue score will be derived at each time point (Baseline, Month 3, 6 and 12) and the change from baseline calculated for Month 3, 6 and 12.

6.1.2.9. Change from Baseline in Morning Stiffness

The duration of morning stiffness is measured in minutes and is collected at all time points (Baseline, Month 3, 6 and 12).

6.1.2.10. Change from Baseline in WPAI

The Work Productivity and Activity Impairment (WPAI, V2.0 15Jan2007) questionnaire consists of 6 questions, from which scores are derived for four domains; absenteeism, presenteeism (impairment at work/reduced job effectiveness), work productivity loss and activity impairment. These are calculated as percentage impairment as follows:

1. Absenteeism = (Q2 / (Q2 + Q4)) *100
2. Presenteeism = (Q5 / 10) * 100
3. Work Production Loss = { ((Q2 / (Q2 + Q4)) + [(1 – (Q2 / (Q2 + Q4))) * (Q5/10)] }* 100
4. Activity Impairment = (Q6/10)* 100

Note that higher numbers indicate greater impairment and less productivity.

A domain score can not be calculated if there is a missing response to the items required to derive the score (http://www.reillyassociates.net/WPAI_Coding.html).

WPAI scores are based on 1-item (presenteeism, activitity impairment), 2-items (absenteeism) and multiple items (overall work productivity); a score can not be calculated if there is a missing response to the corresponding item. An analysis of

absenteeism would include only those subjects with responses to hours missed and to hours worked; an analysis of presenteeism would include only those subjects with a response to productity while working. For the analysis of overall work productivity, scores can only be calculated for subjects with responses to hours missed, hours worked and productivity while working.

See Section 7.1.4 for the handling of missing data in the analysis of WPAI.

The percentage impairment will be derived for each domain at each time point (Baseline, Month 3, 6 and 12) and the change from baseline calculated for all post-baseline timepoints (Month 3, 6 and 12).

6.1.2.11. Patient satisfaction with Treatment

Patient's satisfaction with treatment is assessed at Baseline and the final visit (month 12) on a 5-point scale (where 0 = very dissatisfied and 4 = very satisfied).

6.1.2.12. Patient Global Assessment (PtGA)

Patient Global Assessment (PtGA) is measured using a visual analogue scale (VAS) on a 0-100 mm scale, where 0 mm = no symptoms of disease and 100 mm = maximum seriousness which a patient can imagine. PtGA is measured at each time point (Baseline, Month 3, 6 and 12).

6.1.2.13. Physician Global Assessment (PhGA)

Physician Global Assessment (PhGA) is measured using a VAS on a 0-100 mm scale (where 0 = no disease activity and 100 = maximum disease activity) at each time point (Baseline, Month 3, 6 and 12).

6.2. Safety Endpoints

The reporting of safety data will be in accordance with Pfizer Data Standards.

Safety endpoints in this study include the incidence of all adverse events (AEs) and serious adverse events (SAEs). These will be summarized by the treatment group 'Tofacitinib' (see Section 8.2.2).

The treatment patterns of RA patients prescribed Tofacitinib in a real-world setting will also be described. Patients will be classified into 6 discrete categories as follows:

1. **Tofa Mono**
   Patients who receive Tofacitinib as a monotherapy throughout the study.

2. **Tofa Combo**
   Patients who receive Tofacitinib in combination with another DMARD throughout the study.

3. **Tofa Mono to Combo**
   Patients who start the study receiving Tofacitinib as a monotherapy and then switch to Tofacitinib in combination with another DMARD.

4. **Tofa Combo to Mono**
   Patients who start the study receiving Tofacitinib in combination with another DMARD and then switch to Tofacitinib as a monotherapy.

5. **Tofa Mono to Other Treatment**
   Patients who start the study receiving Tofacitinib as a monotherapy and then discontinue Tofacitinib and switch to another DMARD.

6. **Tofa Combo to Other Treatment**
   Patients who start the study receiving Tofacitinib in combination with another DMARD and then discontinue Tofacitinib.

The duration of treatment with Tofacitinib (in days) will be calculated as:

Date of permanent Tofacitinib discontinuation – Start Date of Tofacitnib treatment + 1

In addition, the time to first treatment change will be calculated:

Date of treatment change – Study Start Date + 1

Where a treatment change is defined as the start of a new treatment or the discontinuation of a current treatment.

6.3. Other endpoints

6.3.1. Healthcare Resources and Costs

The use of healthcare resources and costs in patients with RA treated with Tofacitinib in Greece will be examined using the following:

- The number of visits to a rheumatologist or other RA specialist since the last visit.
- The number of patients with diagnostic tests and the number of diagnostic tests since the last visit. Diagnostic tests include: WBC, ESR/CRP, glucose, hepatic enzymes, Urea, Creatinine, lipid profile, X-rays, other (summarised by visit).
- Hospitalisations since the last visit:
  The number of patients who are hospitalised, the number of hospitalisations per patient and the number of days hospitalised per patient, the reason, overall and by type of hospitalisation (day care or overnight stay) will be derived.(summarised by visit).

In addition an economic evaluation will also be performed. This is not covered in this SAP.

6.4. Covariates

The following potential risk factors will be included as covariates in the logistic regression analyses of remission and LDA:

- Age:
  Age at baseline will be included as a continuous covariate.
- Gender (male/female).
- Weight:
  The weight (kg) at baseline will be included as a continuous covariate.
- Smoking history:
  Smoking history will be included as a categorical covariate with five levels:
- Never;
- History but cessation of more than 3 years;
- History but cessation of less than 3 years; □ Smoker, regular (≥1 cigarette/day); □ Smoker, low (<1 cigarette/day).

If there are issues with the convergence of the model using the above categories, these will be collapsed into the following 3 categories:

- Never;
- Ex-Smoker; □ Smoker.

- Duration of Disease at Baseline:
  Time (years) from RA diagnosis to inclusion in the study will be included as a continuous covariate.

- Serology at Baseline:
- Rheumatoid Factor (RF) will be included with two levels: positive and negative.
- Anti-citrullinated protein antibodies (ACPA) will be included with two levels: positive and negative.

- Previous DMARD use (number):
  The number of previous biological (bDMARDS) will be included as a continuous covariate. See Appendix A1.2 for details of drugs to be considered as bDMARDS.

- Current concomitant treatment:
  Two covariates will be fitted, one for the most common concomitant treatment (Yes/No) and one for the second most common concomitant treatment (Yes/No).

- Glucocorticoid dose:
  The Glucocorticoid dose at baseline will be included as a continous covariate.

- csDMARDs:
- The number of current conventional synthetic DMARDS (csDMARDS) at baseline will be included as a continuous covariate. See Appendix A1.2 for details of drugs to be considered as csDMARDS.
- The dose of csDMARDs at baseline will be included as a continuous covariate.

- Acute phase reactants:
  ESR and CRP at baseline will be included as continous covariates.

- DAS28-4 (CRP) score.

- DAS28-4 (CRP) at baseline will be included as a continuous covariate. This will be calculated as outlined in Section 6.1.1.

- HAQ-DI score:
  The HAQ-DI score at baseline will be included as a continuous covariate. This will be calculated as outlined in Section 6.1.2.3.

- CDAI:
  CDAI at baseline will be included as a continuous covariate. This will be calculated as outlined in Section 6.1.2.4.

- Lipids:

The change from baseline at month 3 in Low-density lipoprotein (LDL) will be included as a continuous covariate.

- Hb levels:
  The change from baseline at month 3 in hemoglobin (Hb) will be included as a continuous covariate.

7. HANDLING OF MISSING VALUES

For the primary endpoints of remission and LDA, a non-responder imputation (NRI) will be used for patients who have all components of DAS28-4 (CRP) missing at Month 6. This means that if a patient has all components of the DAS28 (CRP) missing at Month 6, they will be classified as a non-responder (i.e. not having remisson or LDA respectively) for the purposes of the primary analyses. If a patient has at least one non-missing component of the DAS28-4 (CRP) at Month 6, then missing components will be imputed by carrying forward the last non-missing post-baseline value, before deriving the DAS28-4 (CRP) at Month 6 in order to determine whether they have acheived remission or LDA at Month 6.

For EQ-5D, FACIT-Fatigue Scale and WPAI, missing data will be handled using

the mixed-effect model as described in Section 8.1.2. In addition, the methods for handling missing data when deriving these scores described below will also be applied.

7.1. Patient Reported Outcomes (PROs)

For the questionnaires that will be used in the study (EQ-5D, HAQ-DI, FACIT-Fatigue and WPAI), the rules for replacing missing values suggested by the producers of these questionnaires will be followed. If these rules result in the patient having a missing score at a visit, then the missing values will be assumed to be missing at random. Further details of the missing data rules from the relevant scoring guidelines are provided below.

7.1.1. HAQ-DI

There must be responses in at least 6 of the 8 domains for a HAQ-DI score to be calculated. If there are 2 or more missing domains, the HAQ-DI will be set to missing. In the event that one or two domains have missing data, the score will be calculated by taking the average over the non-missing domains as follows:

HAQ-DI = Sum of non-missing domain scores/number of non-missing domains.

7.1.2. EQ-5D

Missing weighted dimension scores are replaced by the mean of the non-missing weighted dimension scores. If a weighted score is missing and replaced in this way by a mean weighted score of zero, step 3 of the algorithm in Section 6.1.2.7 is applied.

7.1.3. FACIT-Fatigue Scale

If there are missing responses to items, the following approach will be taken: If more than 50% of the items are non-missing, the score will be prorated by multiplying the sum of the subscale by the number of items in the subscale, then dividing by the number of items actually answered. This can be written as:

Prorated FACIT-Fatigue score = [Sum of item scores] **x** 13□□ [N of items answered]

The score is only considered valid if more than 50% of the items are answered. Therefore if 50% or more of the items have a missing response, the FACIT fatigue score will be set to missing.

7.1.4. WPAI

A domain score can only be calculated if none of the required items have missing data. If any of the required items have missing data the domain score will be set to missing. Missing data will not be imputed.

8. STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1. Statistical methods

8.1.1. Logistic Regression Analysis

The potential role of several baseline clinical and biochemical parameters as independent predictive factors of remission or LDA will be investigated using logistic regression. The potential predictive factors to be considered are described in Section 6.4.

The effect of each factor will firstly be tested with a univariate logistic regression model. All potential factors with a p-value resulting from univariate analysis less than 10% will then be included in a multivariate model, along with any factors considered important from a clinical point of view. Predictive factors and their interactions will be chosen using a backward (step-down) selection. Predictive factors chosen will be those which have a p-value less than 5%. In the case of a significant interaction, all individual components of the interaction will be "forced" in the model, whether they are statistically significant or not.

The SAS procedure PROC LOGISTIC will be used to perform the logistic regression analyses.

8.1.2. Analysis of Continuous Data

Unless specified otherwise, continuous variables will be summarised using descriptive statistics (n, mean, standard deviation (SD), median, min, max).

The change from baseline in EQ-5D (Section 6.1.2.7), FACIT-Fatigue Scale (Section 6.1.2.8), WPAI (Section 6.1.2.10) at each time point will be analyzed using mixed models for repeated measures (MMRM), with time-point as a fixed effect; the baseline value will be entered in the model as a covariate, alongside the baseline by time-point interaction. Any missing observations will be set to missing and the MMRM model will assume data are missing at random. Subjects will be fitted as a random effect where an unstructured covariance structure will be used to model the within-subject errors. The Kenward-Roger

approximation will be used to estimate denominator degrees of freedom. If the model fails to converge autoregressive 1 will be fitted instead.

The SAS procedure PROC MIXED will be used to perform the MMRM analysis.

8.1.3. Analysis of Categorical Data

Unless specified otherwise, categorical variables will be summarised using counts and percentages. Percentages will be based on the number of patients with non-missing data at each visit.

8.1.4. Analysis of Binary Data

Unless specified otherwise, binary variables will be summarised using counts and percentages. Percentages will be based on the number of patients with non-missing data at each visit.

8.2. Statistical Analyses

Statistical Analysis Software, SAS (Version 9.1 or above) will be used for all analyses.

8.2.1. BACKGROUND AND DEMOGRAPHIC CHARACTERISTICS AND STUDY DISPOSITION

The Demographic and baseline variables will be summarised:

Continuous: Mean, SD, median, minimum and maximum

Categorical: Frequency and Percentage of population.

This will include age, sex, height, weight, smoking history, smoking status, educational level, social status, severity of disease, medical history (including RA diagnosis, RF and ACPA, presence of erosions on x-rays, presence of extra-articular manifestations [anemia, RA nodules, RA related vasculitis, amyloidosis or other]), vaccination status, hepatitis and tuberculosis status, comorbidities, co-medication and concomitant medication.

The number and percentage of patients at each time point will be provided along with reasons for withdrawal from the study. For patients who have discontinued Tofacitinib treatment but remain in the study information on the treatments prescribed following study treatment (Tofacitinib) will be listed.

These summaries will be based on the Full Analysis Set.

8.2.2. Safety Analyses

Safety data will be presented in tabular and/or graphical format and summarized descriptively, where appropriate. Safety data will be summarized according to Pfizer standards.

Adverse events will be reported according to the 3-tier approach.

Two sets of 'All Causality' and 'Treatment Emergent' adverse events will be summarised. The first set will contain adverse event data for patients while they were taking Tofacitinib. This AE table will summarise AEs under one treatment heading for Tofacitinib (which includes monotherapy and combination therapy). A drug lag of 28 days will be used to assign treatment emergence to Tofacitinib. The AE listings will also list the individual drugs that were being taken at the time of the AE. The second set will include a separate listing of AEs which occurred under another DMARD, following cessation of Tofacitinib. For the cases where the patient continues to be observed for the full duration of the study following discontinuation of Tofacitinib.

NOTE: A lag of 28 days will be applied following cessation of Tofacitinib in the first table. The non-Tofacitinib listing will apply a drug lag of 0 days. Therefore an event which starts on the first day of a new treatment, e.g. Enbrel will be assigned to Tofacitinib in the first table if the AE is within 28 days of Tofacitinib, and also in the listing as the AE started after the initiation of Enbrel. Therefore any event which occurs during a 28 day lag of discontinuing Tofacitinib will appear under Tofacitinib in the first table, and under the new treatment in the listing.

The treatment patterns of RA patients (as defined in Section 6.2) and the duration of Tofacitinib treatment will be summarized to include n and the percentage in each treatment pattern category. The time to first treatment change (as defined in Section 6.2) will also be summarized overall and by treatment pattern.

Safety analyses will use the Safety Analysis Set.

8.2.3. Efficacy/Effectiveness Analyses

Analyses of the efficacy/effectiveness endpoints will be based on the FAS.

8.2.3.1. Analysis of the Primary Endpoints

The primary endpoints of rate of remission and LDA at month 6 will be analysed as outlined in Section 8.1.1. The rate of remission and LDA will also summarised at each visit as outlined in Section 8.1.4. Missing data will be imputed using the methods outlined in Section 7.

8.2.3.2. Analysis of Secondary Endpoints

The rate of remission and LDA at month 12 will be analysed in a similar manner to the primary endpoint.

The absolute value and change from baseline for DAS28-4 (CRP), DAS28-4 (ESR), HAQ-DI and morning stiffness will be summarised at each visit as outlined in Section 8.1.2.

HAQ-DI categories will be summarised at each visit as outlined in Section 8.1.3.

The change from baseline in EQ-5D, FACIT-Fatigue and WPAI scores will be analysed using MMRM models as described in Section 8.1.2. Absolute values and the change from baseline will also be summarised at each visit as outlined in Section 8.1.2. Individual domain scores for these three questionnaires will also be summarised at each visit as outlined in Section 8.1.2.

EQ-5D categories will be summarised at each visit as as outlined in Section 8.1.3.

The rate of CDAI, SDAI, DAS28-4 (CRP) and DAS28-4 (ESR) remission will be summarised at each visit as outlined in Section 8.1.4. The rate of LDA (assessed by CDAI, SDAI, DAS28-4 (CRP) and DAS28-4 (ESR) respectively) and HAQ-DI response will also be summarised as outlined in Section 8.1.4.

Patient satisfaction with treatment will be summarised as outlined in Section 8.1.3. In addition, patient satisfaction will be summarised by the last treatment received (Tofacitinib or other treatment).

PtGA and PhGA will be sumarised at each time-point as outlined in Section 8.1.2.

8.2.4. Other Analyses

8.2.4.1. Analysis of Healthcare Resources and Costs

Analysis of Healthcare Resources and Costs will be based on the Full Analysis Set.

The number of visits (since the last study visit) to a rheumatologist or other RA specialist will be summarised for each time-point .

The number of patients with diagnostic tests and the number of diagnostic tests (since the last study visit) will be summarised at each time-point by type of test (WBC, ESR/CRP, Glucose, Hepatic enzymes, Urea, Creatinine, lipid profule, X-rays, other) and overall.

The number of patients who are hospitalised, the number of hospitalisations (since the last study visit) per patient and the number of days hospitalised per patient will be summarised overall and by type of hospitalisation (day care or overnight stay) for each time-point. The reason for hospitalisation will also be summarised.

The above summaries will all be produced using the methods outlined in Section 8.1.3.

The summaries of healthcare resources will be produced by the following subgroups:

- Education Status;
- Working Status;
- Disease Severity;
- Treatment;
- Co-morbidities; HAQ-DI at baseline; EQ-5D at baseline.

See Section 5.3 for more detail on the subgroups.

8.2.5. Summary of Analyses

Outcome	Analysis Set	Supports Protocol Objective Number	Subgroups	Statistical Method	Covariates/ Strata	Missing Data
Rate of Remission at Month 6	FAS	1	None	Logistic regression	Potential predictive factors outlined in Section 6.4	NRI
Rate of LDA at Month 6	FAS	1	None	Logistic regression	Potential predictive factors outlined in Section 6.4	NRI
Rate of Remission at Month 12	FAS	2	None	Logistic regression	Potential predictive factors outlined in Section 6.4	NRI
Rate of LDA at Month 12	FAS	2	None	Logistic regression	Potential predictive factors outlined in Section 6.4	NRI
EQ-5D	FAS	2	None	MMRM	Baseline score, timepoint fitted as a fixed effect and baseline by timepint interaction	MMRM
FACIT-Fatigue	FAS	2	None	MMRM	Baseline score, timepoint fitted as a fixed effect and baseline by timepint interaction	MMRM
WPAI	FAS	2	None	MMRM	Baseline score, timepoint fitted as a fixed effect and baseline by timepint interaction	MMRM

9. LIST OF TABLES AND TABLE SHELLS

A list of tables (LOT) will be produced and will be provided in a separate document.

10. REFERENCES

1. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. PharmacoEconomics 1993; 4(5):353-65.

2. Reilly Associates [Internet]. WPAI Coding Guidelines; c2002 [cited 2019 Oct 14]. Available from: http://www.reillyassociates.net/WPAI_Coding.html

3. Dolan, Paul & Gudex, Claire & Kind, Paul & Williams, Alan. (1995). A Social Tariff for EuroQol: Results from a UK General Population Survey. University of York: Center for Health Economics. 138.

11. APPENDICES

11.1. APPENDIX 1: DATA DERIVATION DETAILS

A1.1 Definition and use of visit windows in reporting

Visit windows will be used for efficacy variables, and for any safety displays that display by month.

Visit Label	Target Day	Definition [Day window]
Enrolment	1	Day 1 or before
Month 3	91	Day 2 to 137
Month 6	183	Day 138 to Day 274
Month 12	365	Day 275 onwards

If two or more visits fall into the same window, keep the one closest to the Target Day. If two visits are equal distance from the Target Day in absolute value, the later visit should be used.

(Safety analysis may follow Pfizer standard).

A1.2 Definitions of bDMARDS and CSDMARDS

The following drug names will be used to determine whether drugs are bDMARDS or csDMARDS

List of bDMARD (Brand name should be used to be able to track biosimilars)
Humira (Adalimumab)

Cimzia (Certolizumab Pegol)
Enbrel (Etanercept)
Simponi (Golimumab)
Remicade (Infliximab)
Orencia (Abatacept)
Roactemra (Tocilizumab)
Kineret (Anakinra)
Mabthera (Rituximab)
Benepali (Etanercept)
Inflectra (Infliximab)
Other

List of csDMARD

Methotrexate
Leflunomide
Hydroxychloroquine sulfate
Sulfasalazine
Other

CCI