Non-Interventional Study Protocol
A3921332

A Non-Interventional Multinational Study of Tofacitinib in Patients Treated for Psoriatic Arthritis

Statistical Analysis Plan
(SAP)

Version: 2

Author: PPD
,
IQVIA
PPD

Date: 22-Aug-2024

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse Event
AIC	Akaike Information Criteria
BSA	Body Surface Area
BMI	Body Mass Index
CI	Confidence Interval
CRP	C-reactive Protein
DAPSA	Disease Activity in Psoriatic Arthritis
DI	Disability Index
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
ENR	Enrolled Set
ESR	Erythrocyte Sedimentation Rate
EULAR	European League Against Rheumatism
FAS	Full Analysis Set
HAQ	Health Assessment Questionnaire
HRQoL	Health-Related Quality of Life
ICF	Informed Consent Form
LDA	Low Disease Activity
LDI	Leeds Dactylitis Index
LEI	Leeds Enthesitis Index
LS	Least Squares
MAR	Missing at random

Abbreviation	Definition
MCP	Metacarpal Phalangeal
MCS	Mental Component Summary
MDA	Minimum Disease Activity
MI	Multiple Imputation
MTP	Metatarsal Phalangeal
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model for Repeated Measures
MTX	Methotrexate
NI	Non-interventional
NRS	Numerical Rating Scale
NSTEMI	Non-ST-elevation Myocardial Infarction
PASDAS	Psoriatic Arthritis Disease Activity Score
PASI	Psoriasis Area and Severity Index
PCS	Physical Health Component Summary
PIP	Proximal Interphalangeal
PsA	Psoriatic Arthritis
PsAID12	Psoriatic Arthritis Impact of Disease (12 questions)
PRO	Patient Reported Outcome
PT	Preferred Term
QoL	Quality of Life
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan

Abbreviation	Definition
SD	Standard Deviation
SF-36	Short Form 36
SI	International system of units
SJC	Swollen Joints Count
SOC	System Organ Class
SPARCC-EI	Spondyloarthritis Research Consortium of Canada Enthesitis Index
STEMI	ST-elevation Myocardial Infarction
TJC	Tender Joints Count
TLFs	Tables, Listings and Figures
VAS	Visual Analog Scale
VTE	Venous Thromboembolism
WHO	World Health Organization

1 AMENDMENTS FROM PREVIOUS VERSIONS(S)

Version	Issue Date	Change Type (New, Revise, Admin)	Summary of Changes
1	22-Nov-2023	New	Not Applicable
2	22-Aug-2024	Revise	Section 7: Under subheading "Missing or partial treatment end dates will be addressed as follows:", updated the line 'first day prior to the day' as 'first day prior or after to the day'. Section 7: Added section "Methods for multiple imputation of the components of primary composite endpoint". Section 8.1.3: Revised the section by removing the line "The covariance structure converging to the best fit, as determined by Akaike's information criterion, will be used as the primary analysis". Section 8.2.9: In the table, column for **Covariates/ Strata** is updated for different secondary objectives. Section 11.1: Updated the details of derivation of the treatment line. Section 11.2: Updated the definition of visit windowing.

2 INTRODUCTION

This statistical analysis plan (SAP) describes full details of the planned statistical analyses including the rules and conventions to be used in the presentation and analysis of patient characteristics, tofacitinib effectiveness on disease activity, and quality of life (QoL) in patients with psoriatic arthritis (PsA) in 8 countries (Belgium, Denmark, Finland, France, Israel, Netherlands, Spain and Sweden) who were treatment naïve for tofacitinib at the time of enrollment.

This SAP is based on Protocol number A3921332, version 2.0, dated 14-Feb-2023 and electronic Case Report Forms (eCRFs) version 5.0 dated 2-Nov-2023.
Note: in this document any text taken directly from the non-interventional (NI) study protocol is italicized.

2.1 STUDY DESIGN

This is a prospective, non-interventional, multinational, multi-center study to evaluate the effectiveness of tofacitinib in patients with PsA in routine clinical practice with up to 12 months of follow-up. Patients must be treatment naïve for tofacitinib at the time of enrollment.
Patients will be followed up during routine clinic visits, consistent with the European League Against Rheumatism (EULAR) guidelines for follow-up of patients with PsA. Documented visits for the purpose of this study will be at enrollment, then at approximately Months 3 (visit 2), 6 (visit 3), and 12 (visit 4, end of study visit) following enrollment and initiation of treatment with tofacitinib. Patients that permanently discontinue treatment with tofacitinib will be discontinued from the study. Study follow-up will stop at Month 12 unless a patient has already discontinued. No additional diagnostic or monitoring procedures are mandated, and follow-up visits are captured as part of normal medical practice.

Study population

This study aims to enroll up to 200 PsA patients who are treatment naïve for tofacitinib at the time of their enrollment and have decided to initiate treatment with tofacitinib. Please refer to section 9.5 in the protocol for details about the sample size calculations.

This study will enroll approximately 200 patients from rheumatologists and/or PsA specialist centers in 8 countries (Belgium, Denmark, Finland, France, Israel, Netherlands, Spain, and Sweden) over an enrollment period of 25 months. The countries listed above are anticipated to participate in this study, however additional countries may be included at a later time. Each patient will have up to 12 months of follow-up for a total study duration of 37 months. Consecutive patients attending a routine clinical visit will be invited to participate if they meet the eligibility criteria for the study and are to start on treatment with tofacitinib for active PsA.

Data source

Patient data will be entered into an eCRF located on a secure web-based electronic data entry (EDC) system. No visits or examinations, laboratory tests, or procedures are mandated as part of this study. Investigators will be requested to perform commonly used PsA disease assessments. Patients will be requested to complete validated PROs

at each study visit, which will be completed in paper form in local language and entered in the EDC system by IQVIA.

2.2 STUDY OBJECTIVES

The aim of this study is to evaluate the effectiveness of treatment with tofacitinib on disease activity, remission, and QoL in patients with PsA, over a 12-month observation period.

Primary Objective

- *To evaluate the proportion of patients achieving low disease activity (LDA) after 6 months of follow-up.*

Secondary Objectives

- *To evaluate the proportion of patients achieving LDA after 3 and 12 months of follow-up.*
- *To evaluate the proportion of patients achieving minimum disease activity (MDA) after 3, 6, and 12 months of follow-up.*
- *To evaluate the proportion of patients achieving remission after 3, 6, and 12 months of follow-up.*
- *To evaluate change from baseline in Psoriatic Arthritis Impact of Disease 12 Questions (PsAID12) score after 3, 6, and 12 months of follow-up.*
- *To evaluate change from baseline in Spondylarthritis Research Consortium of Canada Enthesitis Index (SPARCC-EI) score after 3, 6, and 12 months of follow-up.*
- *To identify prognostic indicators for achieving LDA after 3, 6, and 12 months of follow-up.*
- *To evaluate change from baseline in QoL scores after 3, 6, and 12 months of follow-up.*

Exploratory Objectives

- *To evaluate changes in measures of disease activity (Psoriatic Arthritis Disease Activity Score [PASDAS] and Disease Activity in Psoriatic Arthritis [DAPSA]) from baseline after 3, 6, and 12 months of follow-up.*

3 INTERIM ANALYSES

No interim analysis is planned as part of this study.

4 HYPOTHESES AND DECISION RULES

No hypothesis testing or comparisons between groups are planned for this study.

5 ANALYSIS SETS/POPULATIONS

Three different analysis sets will be considered for this study, all analysis sets are described below.

5.1 FULL ANALYSIS SET

The full analysis set (FAS) will contain all enrolled patients (see definition in section 5.3.1) who have received at least one dose (including partial dose) of study medication.

5.2 SAFETY ANALYSIS SET

The safety analysis set (SAF) will contain all patients who have received at least one dose (including partial dose) of study medication independently of inclusion/exclusion criteria.

5.3 OTHER ANALYSIS SETS

5.3.1 All Patients Enrolled Analysis Set

The all patients enrolled set (ENR) will contain all patients who provide informed consent and meet all eligibility criteria to participate in this study.

Inclusion criteria:

Patient eligibility should be reviewed by an Investigator before patients are included in the study. Patients must meet all of the following inclusion criteria to be eligible for inclusion in the study:

1. *Patients aged ≥18 years.*
2. *Moderate to severe PsA disease activity diagnosed.*
3. *Patients for whom the physician's decision has been made to initiate treatment with tofacitinib, in usual clinical practice conditions and in compliance with the local label.*
4. *Patients are treatment naïve to tofacitinib on the date of providing informed consent.*
5. *Evidence of a personally signed and dated informed consent document indicating that the patient (or a legally acceptable representative) has been informed of all pertinent aspects of the study.*
6. *Patients on Disease-Modifying Antirheumatic Drug (DMARDs) must have not had a treatment change in the past 3 months.*

Exclusion criteria:

Patients meeting any of the following criteria will not be included in the study:

1. *Contraindications according to the Xeljanz® (tofacitinib) prescribing information.*
2. *Receipt of any investigational drug within 3 months before study inclusion.*
3. *Patient is pregnant or breastfeeding.*
4. *Recent herpes zoster infection (within past 6 months) or history of severe disseminated herpes zoster infection.*
5. *Active treatment for a malignancy.*
6. *Concomitant treatment with biological Disease-Modifying Antirheumatic Drugs (bDMARD).*

5.4 SUBGROUPS

Descriptive analyses will be stratified, where specified, by:

- Body Mass Index (BMI) at enrollment (<18.5, 18.5 to <25, 25 to <30, ≥30 kg/m^2);
- Treatment line (tofacitinib as monotherapy, combination therapy with Methotrexate [MTX], combination with other csDMARD); derived as described in section 11.1)

- Duration of symptoms of at least 6 weeks prior to enrollment, stratified by patients with duration of current episode of symptoms prior to enrollment ≥6 weeks and <6 weeks;
- Erythrocyte sedimentation rate (ESR) at enrollment (Abnormal, Not abnormal, derived as described in section 11.1);
- C-reactive protein (CRP) at enrollment (Abnormal, Not abnormal, derived as described in section 11.1);
- Rheumatoid nodules at enrollment (Present, Absent);
- Unequivocal radiological erosion at enrollment (Present, Absent, derived as described in section 11.1);
- Unequivocal bony decalcification localized to the joints of the hands and wrists at enrollment (Present, Absent);
- Symmetric arthritis at enrollment (No, Yes);
- Arthritis ≥3 joint areas at enrollment (No, Yes, derived as described in section 11.1);
- Arthritis of the hand joints at enrollment (No, Yes, derived as described in section 11.1);
- Arthritis of only 1 medium-large joint at enrollment (No, Yes, derived as described in section 11.1);
- Arthritis of 2-10 medium-large joints and/or 1-3 small joints at enrollment (No, Yes, derived as described in section 11.1);
- Arthritis of 4-10 small joints with or without involvement of large joints at enrollment (No, Yes, derived as described in section 11.1);
- Arthritis of >10 joints at enrollment (with at least one small joint) (No, Yes, derived as described in section 11.1).

6 ENDPOINTS AND COVARIATES

6.1 EFFECTIVENESS ENDPOINTS

6.1.1 Primary Effectiveness Endpoint

The primary effectiveness endpoint is achieving LDA at Month 6. If applicable, the results will be presented stratified by PASDAS status at enrollment (PASDAS ≤3.2, PASDAS >3.2).

Achieving LDA will be defined as: Yes, if PASDAS score ≤3.2; No, if PASDAS score >3.2. The PASDAS is a composite disease activity measure (range 0–10). Details on the definition and derivation of the PASDAS can be seen in the section 11.1.

6.1.2 Secondary Effectiveness Endpoints

Secondary effectiveness endpoints are:

- LDA achievement defined as PASDAS score ≤3.2 at Months 3 and 12. If applicable, the results will be presented stratified by PASDAS status at enrollment (PASDAS ≤3.2, PASDAS >3.2).

- MDA achievement at Months 3, 6, and 12.
  A patient is classified as in MDA when they meet 5 of 7 of the following criteria:
  - Tender joint count [TJC66] ≤1

- Swollen joint count [SJC68] ≤1
- PASI ≤1 or BSA ≤3
- Patient pain assessment Visual Analog Scale (VAS) ≤15
- Patient global assessment VAS ≤20
- Health Assessment Questionnaire- Disability Index (HAQ- DI) ≤0.5
- Tender enthesial points ≤1

More details can be seen in the section 11.1

- Remission achievement defined as PASDAS score ≤1.9 at Months 3, 6, and 12 (see section 11.1).
- Remission achievement defined as DAPSA score ≤4.0 at Months 3, 6, and 12 (see section 11.1).
- Change from baseline in PsAID12 score at Months 3, 6, and 12. (see section 11.1).
- Change from baseline in SPARCC-EI score at Months 3, 6, and 12. (see section 11.1).
- LDA achievement at Months 3, 6, and 12 based on presence of prognostic factors listed in section 0.
- Change from baseline in QoL using PRO scores at Months 3, 6, and 12. (see section 11.1).

6.1.3 Exploratory Effectiveness Endpoints

- Change from baseline in PASDAS scores at Months 3, 6, and 12. (see section 11.1).
- Change from baseline in DAPSA scores at Months 3, 6, and 12. (see section 11.1).
- Change from baseline in ESR scores at Months 3, 6, and 12.
- Change from baseline in CRP scores at Months 3, 6, and 12.

6.2 SAFETY ENDPOINTS

- Adverse Events (AE), serious adverse event (SAE), AE leading to study treatment discontinuation, and AE with a reasonable possibility that the event is related to tofacitinib will be considered.
- Laboratory results will be presented.

6.3 PSORIATIC ARTHRITIS DISEASE ACTIVITY SCORE (PASDAS)COVARIATES

When specified (see sections 8.1.3 and 8.2.2), effectiveness endpoints will be analyzed using suitable regression analyses, with adjustment for the following covariates:

- age,
- prognostic factors specified in section 0.

7 HANDLING OF MISSING VALUES

The number of non-missing observations for all variables will be indicated in descriptive tables.

Partial dates for any clinical event will be handled in the following manner:

- Missing day (only) will be treated as the 15th day of the month unless such imputation is illogical; in the latter case, the date will be set to the first day prior to the day of the event or measurement that makes the imputed day illogical.
- Missing month and/or year will not be imputed.

Missing or partial treatment end dates will be addressed as follows:

- Missing day (only) will be treated as the 15th day of the month unless such imputation is illogical; in the latter case, the date will be set to the first day prior or after to the day of the event or measurement that makes the imputed day illogical.
- Treatment end date is missing then death date or study discontinuation date, will be used (whichever comes first).
- If all the above are missing, then the date of the last observed value from the last visit, but only if treatment discontinuation is indicated at that visit.
- Finally, if the above dates are missing, then the final data cut date of the study will be considered as the treatment end date.

Partial dates not already covered above will be handled in the following manner:

- Missing day (only) will be treated as the 15th day of the month.
- Missing month (only) will be assumed to be July.
- Missing day and month will be assumed to be July 1.
- Missing year will not be imputed.

Any missing component in any composite endpoint will result in the composite endpoint as missing.

The PRO questionnaires will use standard scoring rules, which include any information for the handling of missing data. In the absence of scoring guidelines for missing data, the individual items will be reported as missing and the summary score itself will be considered missing.

Missing data for the primary effectiveness categorical endpoint

- All missing data consecutive to a study discontinuation will be considered as a Non-Response.
- All other missing data (especially in case of non-monotone patterns) will be replaced using multiple imputation under a missing at random (MAR) assumption.

- For the components of the primary composite endpoint (LDA at baseline and at 6 months) published scoring algorithms and missing values, rules will be used before any multiple imputation (MI), if available.
  Missing data (components of the primary composite endpoint at enrollment and at 6 months) will be imputed using SAS PROC MI under the assumption that the missing values are MAR. This assumption requires that missing values can be correlated to measured covariates, but not with unmeasured covariates. This MAR assumption underlying MI is plausible. The imputation model will include at least all prognostic factors listed in section 0 as well as any additional variables that thought to influence missingness. At least 25 imputation datasets will be generated, and

results will be aggregated across these imputed datasets using SAS PROC MI ANALYZE.

Methods for multiple imputation of the components of primary composite endpoint:

- **Logistic regression** will be used for binary and ordinal variables.
- **Discriminant analysis** will be used for nominal variables.
- **Predictive mean matching** (PMM) will be used for continuous and count variables. The method involves prediction, matching, and imputation steps to replace missing values with observed values that are close to the predicted value, and thus, it preserves the distribution of the observed data.

    1. Prediction: For each variable with missing values, a regression model is fitted using the other variables in the dataset as predictors. This model is used to predict the missing values.
    2. Matching: Instead of just filling in the predicted value, PMM finds observed values that are close to the predicted value[1]. It then randomly selects one of these observed values to use as the imputed value.
    3. Imputation: The selected observed value is used to replace the missing value.

    This process is repeated for each variable with missing values, and then multiple times to create multiple imputed datasets[1].
    The advantage of PMM is that it preserves the distribution of the observed data because the imputed values are always observed values. This makes it a flexible and robust method for multiple imputation

- Missing data for the secondary effectiveness categorical endpoints:
  All missing data will be considered as a Non-Response and will not be imputed

8 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1 STATISTICAL METHODS

8.1.1 Statistical Methods Analyses for Continuous Data

Unless otherwise specified, continuous and discrete data will be summarized by time point in tables containing descriptive statistics: mean, standard deviation (SD), median, first and third quartiles (Q1-Q3), range (minimum-maximum) and the number of non-missing and missing observations.

Mean, SD, median, Q1 and Q3 will be presented with 1 more decimal than the raw data, the minimum and maximum will be presented with the same number of decimals as for the original measurement.

For quantitative measurements, change from baseline will be calculated as: Value at Visit X – Baseline Value and displayed with the same number of decimals as for the original measurement.

8.1.2 Analyses for Categorical Variables

Categorical variables (including binary variables) will be summarized as the number of patients and percentages (%) of patients in each category. Percentages will not include the missing category and will be calculated over the number of patients with available (non-missing) data. The count of missing observations and the number of patients in each category will be provided in all tables.

Computed percentages will be presented with 1 decimal place. Percentages equal to 100 will be presented as 100% and no percentage will be presented for zero frequencies.

When applicable, 95% confidence intervals (CI) for the proportions will be calculated using the Agresti-Coull method.

8.1.3 Mixed Model for Repeated Measures

Changes from baseline at specified follow-up assessment time points will be analyzed using MMRM model under the MAR framework. The models will include visit (including baseline) as categorical fixed effect factor. An unstructured covariance structure will be assumed for the within-subject repeated measures. If models do not converge, a heterogeneous Toeplitz covariance structure followed by the heterogeneous first-order autoregressive covariance structure will be considered. Parameters will be estimated using the Newton-Raphson algorithm. Denominator degrees of freedom will be estimated using the Kenward-Roger approximation.

The potential prognostic factors defined in the section 0 and other patient baseline characteristics that are known potential confounders and were not identified by this procedure may be included in the final model, if agreed by the team.
Least squares (LS) means at each time point and for change from baseline, standard errors and 95% CI will be presented by visit in the framework of this model. Associated p-values will be rounded to 3 decimal places.

8.1.4 Logistic Regression

Logistic regression will be used to identify prognostic factors for achieving LDA after 3, 6, and 12 months among eligible patients who initially exhibited no LDAA univariate logistic regression will be run for each of the potential prognostic factors. If the variable significantly impacts LDA at the 10% level then it will be put forward into a multivariable model provided that the sample size is deemed adequate according to the events per predictor rule of ten, which necessities that there should be a minimum of ten participants in the least category per predictor.

If the multivariable analysis is conducted, backwards selection will be used to remove factors from the multivariate model which do not have a significant p-value at the 5% level. The covariate with the highest p-value will be removed in turn until all factors have a p-value less than 0.05. A significant interaction is tested for between all variables in the multivariate model if the sample size permits. If a significant interaction is found then the individual factors of this interaction will be kept in the model, regardless of their p-value.

Furthermore, a table showing all the two-sided 95% CI and p-values for the "raw" odds ratios (ORs) coming from the univariate models will be created. This will be for all the factors, regardless of whether they are included in the multivariable model. If the multivariable model is conducted, a similar table will be provided for all the selected predictors. No adjustment for multiplicity will be used.

8.2 STATISTICAL ANALYSES

8.2.1 Safety Analyses

All safety analyses will be performed using the SAF population.

8.2.1.1 Exposure to Treatment

The following measurements will be presented and summarized by means of descriptive statistics. The number and percentage of patients with values in each category will be produced. All variables will be presented by time point and will be reported either as per the eCRF or will be part of derived variables.

- Daily dose of the initial treatment (mg)
  - As a continuous variable.
  - As a categorical variable and presented as: 5mg, 10mg, 11mg, 20mg, and 22mg.
- Frequency of the initial treatment
  - As a categorical variable as: Daily, Twice a day, Other.
- Treatment duration of the initial treatment (months)
  - Will be derived as described in section 11.1 and presented as continuous variable.
- Time from enrollment (signature on Informed Consent Form (ICF)) to first dose (days).
  - Will be derived as described in section 11.1 and presented as continuous variable.
- Action taken
  - As a categorical variable as: Initial dose, Dose reduced, Dose increased, Drug interrupted, Drug withdrawn.
- Reason for dose change/interruption/withdrawal
  - As a categorical variable as: AE, SAE, Lack of Efficacy, Patient's decision, Investigator's decision, Other.

8.2.1.2 **Adverse Events**

Adverse Events, SAE, Non-Serious AE, Non-Serious AE in ≥2% of patients, AE leading to study treatment discontinuation, and AE with a reasonable possibility that the event is related to tofacitinib will be summarized by system organ class (SOC) and preferred term (PT) using the Medical Dictionary for Regulatory Activities (MedDRA) version 23.1.

Adverse Events, SAE, AE leading to study treatment discontinuation, and AE with a reasonable possibility that the event is related to tofacitinib will be identified from the AE page of the eCRF.

Listings will be provided for:

- Cardiovascular events

- Venous thromboembolism (VTE)
- Opportunistic infections
- Cancer
- Death

8.2.1.3 Laboratory Evaluations

Results from the laboratories, as collected by the eCRF, will be included in the reporting of this study for hematology, chemistry, and inflammatory markers.

All relevant clinical laboratory tests will be classified according to investigator assessment as clinically significant (Yes, No).

Quantitative laboratory tests (hematology, chemistry, and inflammatory markers) will be presented using SI Units and will be summarized by means of descriptive statistics. The number and percentage of patients with values in each category of the clinically significant variable will be produced. All the results will be presented by time point.

8.2.2 Analyses of the Effectiveness Endpoints

The following analyses will be performed in the FAS and results will be presented both in overall and by stratified by treatment line (tofacitinib as monotherapy, combination therapy with Methotrexate [MTX], combination with other csDMARD);

8.2.2.1 Primary analysis

- Number and proportion of patients achieving LDA (defined as PASDAS score ≤3.2) at Month 6 will be presented in a summary table together with correspondent baseline distribution, for observed and imputed data. A 95% CI will be provided for the proportion of patients achieving LDA.

8.2.2.2 Secondary analysis

- Number and proportion of patients achieving LDA defined as PASDAS score ≤3.2 at Months 3 and 12 will be presented in a table together with correspondent baseline distribution. A 95% CI will be provided for the proportion of patients achieving LDA.

- Number and proportion of patients achieving MDA defined as at least 5 of 7 criteria met at Months 3, 6, and 12 will be displayed in a table together with the description with the correspondent baseline distribution. A 95% CI will be provided for the proportion of patients achieving MDA.

- Number and proportion of patients achieving remission defined as PASDAS score ≤1.9 at Months 3, 6, and 12 will be presented in a summary table. A 95% CI will be provided for the proportion of patients achieving remission defined as PASDAS score ≤1.9.

- Number and proportion of patients achieving remission defined as DAPSA score ≤4.0 at Months 3, 6, and 12 will be presented in a table together with the

correspondent baseline distribution. A 95% CI will be provided for the proportion of patients achieving remission defined as DAPSA score ≤4.0.

- Change from baseline in the PsAID12 score at Months 3, 6, and 12 will be modeled as defined in section 8.1.3. LS means for change from baseline PsAID12 score with standard errors and 95% CIs will be presented at Months 3, 6, and 12 in a table. If applicable, multivariable models will be also presented.

- Change from baseline in the SPARCC-EI score at Months 3, 6, and 12 will be modeled as defined in section 8.1.3. LS means for change from baseline SPARCC-EI score with standard errors and 95% CIs will be presented at Months 3, 6, and 12 in a table. If applicable, multivariable models will be also presented.

- Number and proportion of patients achieving LDA at Months 3, 6 and 12 will be presented stratified by each one of the potential prognostic factors described in section 0. Univariate and/or multivariable logistic regressions will be used to identify prognostic factors for achieving LDA.

- Change from baseline in the Short Form 36 (SF-36) scores (Physical Health Component Summary Score and the Mental Health Component Summary Score) at Months 3, 6, and 12 will be modeled as defined in section 8.1.3. LS means for change from baseline SF-36 scores with standard errors and 95% CIs will be presented at Months 3, 6, and 12 in a table. If applicable, multivariable models will be also presented.

- Change from baseline in the HAQ-DI score at Months 3, 6, and 12 will be modeled as defined in section 8.1.3. LS means for change from baseline HAQ-DI score with standard errors and 95% CIs will be presented at Months 3, 6, and 12 in a table. If applicable, multivariable models will be also presented.

8.2.2.3 Exploratory analysis

- Change from baseline in the PASDAS score at Months 3, 6, and 12 will be modeled as defined in section 8.1.3. LS means for change from baseline PASDAS score with p-values and 95% CIs will be presented at Months 3, 6, and 12 in a table. If applicable, multivariable models will be also presented.

- Change from baseline in the DAPSA score at Months 3, 6, and 12 will be modeled as defined in section 8.1.3. LS means for change from baseline DAPSA score with standard errors and 95% CIs will be presented at Months 3, 6, and 12 in a table. If applicable, multivariable models will be also presented.

8.2.2.4 Disease Activity Measures and Patient Reported Outcomes

The following disease activity measures and patient reported outcomes will be presented and summarized by means of descriptive statistics. The number and percentage of patients with values in each category will be produced. All variables will be presented by time point and will be reported either as per the eCRF or will be part of derived variables. A summary table with all the disease activity measures will be presented separately and presented by timepoint.

Swollen & Tender/Painful Joint Counts

- Swollen joints
  - Temporomandibular joint
    - As a discrete variable.
  - Sternoclavicular joint
    - As a discrete variable.
  - Acromioclavicular joint
    - As a discrete variable.
  - Glenohumeral(s)
    - As a discrete variable.
  - Elbow(s)
    - As a discrete variable.
  - Wrist(s)
    - As a discrete variable.
  - Metacarpal phalangeal joint
    - As a discrete variable
  - Finger Proximal interphalangeal joint
    - As a discrete variable.
  - Finger Distal interphalangeal joint
    - As a discrete variable.
  - Knees
    - As a discrete variable.
  - Ankle(s)
    - As a discrete variable.
  - Tarsus/Midfoot (feet) (s)
    - As a discrete variable.
  - Metatarsal phalangeal joints
    - As a discrete variable.
  - Toe Proximal Interphalangeal (PIP[s])
    - As a discrete variable.
  - Total swollen joints
    - As a discrete variable.
    - As a categorical variable and presented as: 0, 1-4, >4 swollen joints.

- Tender/Painful joints
  - Temporomandibular joint
    - As a discrete variable.
  - Sternoclavicular joint
    - As a discrete variable.
  - Acromioclavicular joint
    - As a discrete variable.
  - Glenohumeral(s)
    - As a discrete variable.
  - Elbow(s)
    - As a discrete variable.
  - Wrist(s)
    - As a discrete variable.
  - Metacarpal phalangeal joint

- As a discrete variable
- Finger Proximal interphalangeal joint
  - As a discrete variable.
- Finger Distal interphalangeal joint
  - As a discrete variable.
- Hips
  - As a discrete variable.
- Knees
  - As a discrete variable.
- Ankle(s)
  - As a discrete variable.
- Tarsus/Midfoot (feet) (s)
  - As a discrete variable.
- Metatarsal phalangeal joints
  - As a discrete variable.
- Toe PIP(s)
  - As a discrete variable.
- Total tender/painful joints
  - As a discrete variable.
  - As a categorical variable and presented as: 0, 1-4, > 4 tender/painful joints.

Assessment of Disease Activity

- Physician Global Assessment score (mm)
  - As a continuous variable.
- Patient Global Assessment score (mm)
  - As a continuous variable.
- Patient Pain Assessment score (mm)
  - As a continuous variable

Leeds Enthesitis Index (LEI)

- Achilles Tendon Insertions (right and left)
  - As a categorical variable and presented as: Present, Absent.
- Medial Femoral Condyles (right and left)
  - As a categorical variable and presented as: Present, Absent.
- Lateral Epicondyles of the Humerus (right and left)
  - As a categorical variable and presented as: Present, Absent.
- LEI overall score

  Will be derived as described in Section 11.1 and presented as continuous variable.

Leeds Dactylitis Index (LDI)

- Number of dactylitis
  - As a discrete variable [0-20].
  - As a categorical variable and presented as: 0, 1-5, 6-10, 11-15 and 16-20
- Number of fingers and toes with no tenderness
  - As a discrete variable [0-20].

Tender enthesial points

- As a discrete variable.

Psoriasis Area and Severity Index (PASI)

- Head
  - Plaque Characteristics
    - Erythema
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Induration
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Desquamation
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
  - Percentage Area Affected
    - As a categorical variable and presented as: 0%, 1% - 9%, 10% - 29%, 30% - 49%, 50% - 69%, 70% - 89%, 90% - 100%)
  - Lesion score sum
    - Will be derived as described in section 11.1 and presented as continuous variable.
- Upper Limb
  - Plaque Characteristics
    - Erythema
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Induration
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Desquamation
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
  - Percentage Area Affected
    - As a categorical variable and presented as: 0%, 1% - 9%, 10% - 29%, 30% - 49%, 50% - 69%, 70% - 89%, 90% - 100%)
  - Lesion score sum
    - Will be derived as described in section 11.1 and presented as continuous variable
- Trunk
  - Plaque Characteristics
    - Erythema
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Induration
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Desquamation

- As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
- Percentage Area Affected
  - As a categorical variable and presented as: 0%, 1% - 9%, 10% - 29%, 30% - 49%, 50% - 69%, 70% - 89%, 90% - 100%)
- Lesion score sum
  - Will be derived as described in section 11.1 and presented as continuous variable.

- Lower Limbs
  - Plaque Characteristics
    - Erythema
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Induration
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
    - Desquamation
      - As a categorical variable and presented as: None, Slight, Moderate, Severe, Very Severe.
  - Percentage Area Affected
    - As a categorical variable and presented as: 0%, 1% - 9%, 10% - 29%, 30% - 49%, 50% - 69%, 70% - 89%, 90% - 100%
- PASI score
  - Will be derived as described in section 11.1 and presented as continuous variable.

Body Surface Area (BSA)

- The number of a patient's hand areas affected.
  - As a discrete variable.
- Percentage of body area affected.
  - BSA percentages as a categorical variable and presented as: mild (0%–5%), moderate (>5%–10%), severe (>10%–15%) and very severe (>15%).

Nail Psoriasis

- Will be reported as per the eCRF as a categorical variable including the following categories: No, Yes.

Spondyloarthritis Research Consortium of Canada Enthesitis Index (SPARCC-EI)

- Greater trochanter (right and left)
  - As a categorical variable and presented as: Non-tender, Tender.
- Quadriceps tendon insertion into the patella (right and left)
  - As a categorical variable and presented as: Non-tender, Tender.
- Patellar ligament insertion into the patella and tibial tuberosity (right and left)
  - As a categorical variable and presented as: Non-tender, Tender.
- Achilles tendon insertion (right and left)
  - As a categorical variable and presented as: Non-tender, Tender.
- Plantar fascia insertion (right and left)

- As a categorical variable and presented as: Non-tender, Tender.
- Medial epicondyles (right and left)
  - As a categorical variable and presented as: Non-tender, Tender.
- Lateral epicondyles (right and left)
  - As a categorical variable and presented as: Non-tender, Tender.
- Supraspinatus insertion (right and left)
  - As a categorical variable and presented as: Non-tender, Tender.
- SPARCC-EI score

  Will be derived as described in section 11.1 and presented as continuous variable.

Radiological reports scored using Sharp-Van der Heijde modified scoring for PsA.

- Erosion
  - As a continuous variable.
- Joint Space Narrowing
  - As a continuous variable.
- Total Score
  - As a continuous variable.

36 -Item Short Form Health Survey questionnaire

- Physical functioning
  - As a continuous variable.
- Role physical
  - As a continuous variable.
- Bodily pain
  - As a continuous variable.
- General health
  - As a continuous variable.
- Social functioning
  - As a continuous variable.
- Role-emotional
  - As a continuous variable.
- Mental health
  - As a continuous variable.
- Reported health transition score.

  As a categorical variable

Health Assessment Questionnaire (HAQ) – Disability Index

- Disability Index
  - As a continuous variable.
- VAS for pain
  - As a continuous variable.
- Dressing & Grooming
  - As a continuous variable.
- Arising
  - As a continuous variable.
- Eating
  - As a continuous variable.
- Walking
  - As a continuous variable.

- Hygiene
  - As a continuous variable.
- Reach
  - As a continuous variable.
- Grip
  - As a continuous variable.
- Activities
  - As a continuous variable.

Psoriatic Arthritis Impact of Disease (PsAID12)

- Total score
  - As a continuous variable.
- Pain
  - As a continuous variable.
- Fatigue
  - As a continuous variable.
- Skin problems
  - As a continuous variable.
- Work and/or leisure activities
  - As a continuous variable.
- Functional capacity
  - As a continuous variable.
- Discomfort
  - As a continuous variable.
- Sleep disturbance
  - As a continuous variable.
- Coping
  - As a continuous variable.
- Anxiety, fear, and uncertainty
  - As a continuous variable.
- Embarrassment and/or shame
  - As a continuous variable.
- Social participation
  - As a continuous variable.
- Depression
  - As a continuous variable.

8.2.3 Baseline and Other Summaries and Analyses

All the following characteristics will be descriptively described as defined in section 8.1. Protocol deviations and patients' disposition will be presented for all patients in the ENR. All the other characteristic described below will be presented for all patients in the FAS.

8.2.3.1 Patients' disposition

Patient disposition, withdrawals, including inclusion and exclusion criteria will be presented for all patients in the ENR set.

- Patients enrolled (ICF signed)
  - Enrollment failures
  - Patients who did not meet inclusion /exclusion criteria

- Patients in the FAS
  - Patients completing the study
    - Patients completing the study per protocol. Will be reported as per the eCRF as a categorical variable including the following categories: No, Yes.
  - Patients ongoing
  - Patients discontinuing
  - Primary *reason* for discontinuation will be reported as the eCRF as a categorical variable including the following categories: Death, Lost to Follow-Up, Withdrawal by Subject, Site Terminated by Sponsor, Study Terminated by Sponsor, Drug withdrawn, Other.

8.2.3.2 Baseline demographic and disease characteristics

Baseline demographic characteristics will be reported for all patients in the FAS and reported either as per the eCRF or as part of derived variables.

- Age (years)
  - As a continuous variable.
- Sex
  - As a categorical variable and presented as: Male, Female.
- Country
  - As a categorical variable: Belgium, Denmark, Finland, France, Israel, Netherlands, Spain and Sweden.
- PsA Disease History
  - PsA disease duration
    - Will be derived as described in section 11.1 and presented as continuous variable in years
  - Duration of current episode of symptoms prior to enrollment
    - As a continuous variable and categorized and presented as: < 6, ≥6 weeks.
- Current Disease Status
  - Baseline presence of rheumatoid nodules
    - As a categorical variable and presented as: Present, Absent, Not Available.
  - Unequivocal bony decalcification in hand and wrist joints
    - Will be reported as per the eCRF as a categorical variable including the following categories: Present, Absent, Not Available.

8.2.3.3 Baseline potential prognostic factors

All the variables described in section 0 will be summarized in a separate table.

8.2.4 Medical History

Medical conditions will be coded using the MedDRA version 23.1. Variables will be reported either as per the eCRF or will be part of derived variables.

- Condition
  - Will be presented by SOC and PT.
  - Conditions ongoing at enrollment will be stratified according to control status.

8.2.5 Comorbidities

Comorbidities will be presented at baseline and at Month 6. Use of combined hormonal contraceptives and hormone replacement therapy will be described in the Prior and Current Concomitant/PsA Medications subsection.

- Venous thromboembolism (VTE)
  - As a categorical variable and presented as: No, Yes.
    - If yes, type of event, as a categorical variable and presented as: Deep venous thrombosis, Pulmonary embolism.
- Myocardial infarction within previous 3 months
  - As a categorical variable and presented as: No, Yes.
- Myocardial infarction
  - As a categorical variable and presented as: No, Yes.
    - If yes, type of event, as a categorical variable and presented as: ST-elevation myocardial infarction (STEMI), Non-ST-elevation myocardial infarction (NSTEMI), Pulmonary embolism.
- Heart Failure
  - As a categorical variable and presented as: No, Yes.
- Diabetes
  - As a categorical variable and presented as: No, Yes.
    - If yes, type of event, as a categorical variable and presented as: Type I, Type II.
- Hypertension
  - As a categorical variable and presented as: No, Yes.
    - If yes, treatment status and resistance, as a categorical variable and presented as:
      - Treated: No, Yes, Unknown.
      - Treatment resistant: No, Yes, Unknown.
- Inherited coagulation disorder
  - As a categorical variable and presented as: No, Yes.
  - Inherited coagulation disorders will be presented by SOC and PT as described in the Medical History subsection.
- Malignancy
  - As a categorical variable and presented as: No, Yes.
  - Malignancy conditions will be presented by SOC and PT as described in the Medical History subsection.
- Symmetric arthritis
  - As a categorical variable and presented as: No, Yes.

8.2.6 Extra Articular Manifestations

The following measurements will be presented and summarized by means of descriptive statistics. The number and percentage of patients with values in each category will be produced. All variables will be presented by time point and will be reported as per the eCRF.

- Cutaneous involvement
  - As a categorical variable and presented as: Present, Previously present, Never, Not Available.

- Bowel involvement
  - As a categorical variable and presented as: Present, Previously present, Never, Not Available.
- Ocular involvement
  - As a categorical variable and presented as: Present, Previously present, Never, Not Available.
- Cardiovascular involvement
  - As a categorical variable and presented as: Present, Previously present, Never, Not Available.
- Urogenital involvement
  - As a categorical variable and presented as: Present, Previously present, Never, Not Available.
- Pulmonary involvement
  - As a categorical variable and presented as: Present, Previously present, Never, Not Available.

8.2.7 Physical Exam and Smoking Status

The following measurements will be presented and summarized by means of descriptive statistics. The number and percentage of patients with values in each category will be produced. All variables will be presented by time point and will be reported either as per the eCRF or will be part of derived variables.

- Height (cm)
  - Height will be presented in cm. Derivations needed are described in section 11.1.
  - As a continuous variable.
- Weight (kg)
  - Weight will be presented in kg. Derivations needed are described in section 11.1.
  - As a continuous variable.
- Waist Circumference
  - As a continuous variable and presented in cm. Derivations needed are described in section 11.1.
  - As a categorical variable and presented as:
    - for men: ≤94 cm, >94 cm
    - for women ≤80 cm, >80 cm
- BMI
  - Will be derived as described in section 11.1 and presented as categorical variable and presented as: <18.5, 18.5 to <25, 25 to <30, ≥30 kg/m^2.
- Smoking status
  - As a categorical variable and presented as: Never, Current, Quit ≤ 1 year ago, Quit > 1year ago and Unknown.

8.2.8 Prior and Current Concomitant/PsA Medications

Any medications will be coded using the World Health Organization-Drug Dictionary (WHO Drug Q3 2020).
Variables will be reported either as per the eCRF or will be part of derived variables.

PSA treatment was classified into prior or current as described in section 11.1. Percentages will be calculated using the number of patients with non-missing values in the relevant category as a denominator.

- History of prior PsA treatment
  - As a categorical variable as collected in the Demographics & Disease History eCRF page as: No, Yes.
  - If yes,
    - Will be derived and presented as a categorical variable as: 1 prior therapy, 2 prior therapies, 3 prior therapies, 4+ prior therapies.
    - Will be presented by Medication Class / Standardized Medical Term.
- Current PsA treatment
  - As a categorical variable as: No, Yes.
  - If yes,
    - Will be derived and presented as a categorical variable as: Tofacitinib monotherapy, Combined tofacitinib and MTX therapy, Combined tofacitinib and other csDMARD
    - Will be presented by Medication Class / Standardized Medical Term
- Hormonal contraceptive concomitant medication
  - Will be presented by Medication Class / Standardized Medical Term
- AE related concomitant medication.
  - Will be presented by Medication Class / Standardized Medical Term

8.2.9 Summary of Effectiveness and Safety Analyses

Objective	Endpoint	Analysis Set	Subgroups [1]	Statistical Method	Covariates/ Strata	Missing Data
Primary objective: To evaluate the proportion of patients achieving low disease activity (LDA) after 6 months of follow-up	Achieving LDA (PASDAS score ≤3.2) at Month 6	FAS	LDA status at baseline	Descriptive Statistics Count (proportion)	NA	Multiple Imputation
Secondary objective: To evaluate the proportion of patients achieving LDA after 3 and 12 months of follow-up	Achieving LDA (PASDAS score ≤3.2) at Month 3 and 12	FAS	LDA status at baseline	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate the proportion of patients achieving minimum disease activity (MDA) after 3, 6, and 12 months of follow-up	Achieving MDA at Month 3, 6, and 12	FAS	MDA status at baseline	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate the proportion of patients achieving remission after 3, 6, and 12 months of follow-up	Achieving remission (PASDAS score ≤1.9) at Month 3, 6, and 12	FAS	None	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate the proportion of patients achieving remission after 3, 6, and 12 months of follow-up	Achieving remission (DAPSA score ≤4.0) at Month 3, 6, and 12	FAS	None	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate change from baseline in Psoriatic Arthritis Impact of Disease 12 Questions (PsAID12) score after 3, 6, and 12 months of follow-up	Change from Baseline to Month 3, 6, and 12 in PsAID12 score	FAS	None	MMRM	Baseline values of the PsAID12, visit, age, and any other variables identified previously	Observed Data
Secondary objective: To evaluate change from baseline in Spondyloarthritis Research Consortium of Canada Enthesitis Index (SPARCC-EI) score after 3, 6, and 12 months of follow-up	Change from Baseline to Month 3, 6, and 12 in SPARCC-EI score	FAS	None	MMRM	Baseline values of the SPARCC-EI, visit, age, and any other variables	Observed Data

Objective	Endpoint	Analysis Set	Subgroups [1]	Statistical Method	Covariates/ Strata	Missing Data
					identified previously	
Secondary objective: To identify prognostic indicators for achieving LDA after 3, 6, and 12 months of follow-up	Achieving LDA at Months 3, 6 and 12 based on presence of prognostic factors	FAS	BMI, treatment line, Duration of symptoms, ESR, CRP, rheumatoid nodules, unequivocal radiological erosion, unequivocal bony decalcification, symmetric arthritis, arthritis ≥ 3 joint areas, arthritis of the hand joints, arthritis of only 1 medium-large joint, arthritis of 2-10 medium-large joints and/or 1-3 small joints, arthritis of 4-10 small joints with or without involvement of large joints and arthritis of >10 joints at enrollment (with at least one small joint)	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To identify prognostic indicators for achieving LDA after 3, 6, and 12 months of follow-up	Achieving LDA at Months 3, 6 and 12 based on presence of prognostic factors	FAS	None	Logistic Regression	Age, and any other variables	Observed Data

Objective	Endpoint	Analysis Set	Subgroups [1]	Statistical Method	Covariates/ Strata	Missing Data
					identified previously	
Secondary objective: To evaluate change from baseline in Quality of Life (QoL) scores after 3, 6, and 12 months of follow-up	Change from Baseline to Month 3, 6, and 12 in SF-36 physical health component summary score	FAS	None	MMRM	NA	Observed Data
Secondary objective: To evaluate change from baseline in QoL scores after 3, 6, and 12 months of follow-up	Change from Baseline to Month 3, 6, and 12 in SF-36 mental health component summary score	FAS	None	MMRM	NA	Observed Data
Secondary objective: To evaluate change from baseline in QoL scores after 3, 6, and 12 months of follow-up	Change from Baseline to Month 3, 6, and 12 in HAQ-DI score	FAS	None	MMRM	NA	Observed Data
Exploratory objective: To evaluate change from baseline in PASDAS (Psoriatic Arthritis Disease Activity Score) score after 3, 6, and 12 months of follow-up	Change from Baseline to Month 3, 6, and 12 in PASDAS score	FAS	None	MMRM	NA	Observed Data
Exploratory objective: To evaluate change from baseline in DAPSA score after 3, 6, and 12 months of follow-up	Change from Baseline to Month 3, 6, and 12 in DAPSA score	FAS	None	MMRM	NA	Observed Data
Secondary objective: To evaluate the safety of treatment with Tofacitinib	Developing any Adverse Event (AE)	SAF	None	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate the safety of treatment with Tofacitinib	Developing any Serious AE	SAF	None	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate the safety of treatment with Tofacitinib	Developing any Non-Serious AE	SAF	None	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate the safety of treatment with Tofacitinib	Developing any Non-Serious AE in ≥2% of patients	SAF	None	Descriptive Statistics Count (proportion)	NA	Observed Data
Secondary objective: To evaluate the safety of treatment with Tofacitinib	Developing any AE leading to study treatment	SAF	None	Descriptive Statistics	NA	Observed Data

Objective	Endpoint	Analysis Set	Subgroups [1]	Statistical Method	Covariates/ Strata	Missing Data
				Count (proportion)		
Secondary objective: To evaluate the safety of treatment with Tofacitinib	Developing any AE with a reasonable possibility that the event is related to Tofacitinib	SAF	None	Descriptive Statistics Count (proportion)	NA	Observed Data

AE: Adverse Event; BMI: Body Mass Index; CRP: C-reactive Protein; DAPSA: Disease Activity in Psoriatic Arthritis; ESR: Erythrocyte Sedimentation Rate; FAS: Full Analysis Set; HAQ-DI: Health Assessment Questionnaire – Disability Index; LDA: Low Disease Activity; MDA: Minimum Disease Activity; MMRM: Mixed Model for Repeated Measures; NA: Not Available: PASDAS: Psoriatic Arthritis Disease Activity Score; PsAID12: Psoriatic Arthritis Impact of Disease 12 Questions; SAF: Safety Analysis Set; SF-36: Short Form 36; SPARCC-EI: Spondyloarthritis Research Consortium of Canada Enthesitis Index.
[1] All effectiveness analysis will be presented both in overall and by stratified by treatment line (tofacitinib as monotherapy, combination therapy with Methotrexate [MTX], combination with other csDMARD);

9 LIST OF TABLES AND TABLE SHELLS

To be presented in a separate document.

10 REFERENCES

1. Helliwell PS, FitzGerald O, Fransen J, Gladman DD, Kreuger GG, Callis-Duffin K, et al. The development of candidate composite disease activity and responder indices for psoriatic arthritis (GRACE project). Annals of the Rheumatic Diseases. 2013;72(6):986-91.
2. Mease PJ. Measures of psoriatic arthritis: Tender and Swollen Joint Assessment, Psoriasis Area and Severity Index (PASI), Nail Psoriasis Severity Index (NAPSI), Modified Nail Psoriasis Severity Index (mNAPSI), Mander/Newcastle Enthesitis Index (MEI), Leeds Enthesitis Index (LEI), Spondyloarthritis Research Consortium of Canada (SPARCC), Maastricht Ankylosing Spondylitis Enthesis Score (MASES), Leeds Dactylitis Index (LDI), Patient Global for Psoriatic Arthritis, Dermatology Life Quality Index (DLQI), Psoriatic Arthritis Quality of Life (PsAQOL), Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F), Psoriatic Arthritis Response Criteria (PsARC), Psoriatic Arthritis Joint Activity Index (PsAJAI), Disease Activity in Psoriatic Arthritis (DAPSA), and Composite Psoriatic Disease Activity Index (CPDAI). Arthritis Care & Research. 2011;63(S11):S64-S85.
3. Schoels MM, Aletaha D, Alasti F, Smolen JS. Disease activity in psoriatic arthritis (PsA): defining remission and treatment success using the DAPSA score. Annals of the Rheumatic Diseases. 2016;75(5):811-8.
4. Fries JF. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol. 1982;9:789-93.
5. Gossec L, de Wit M, Kiltz U, Braun J, Kalyoncu U, Scrivo R, et al. A patient-derived and patient-reported outcome measure for assessing psoriatic arthritis: elaboration and preliminary validation of the Psoriatic Arthritis Impact of Disease (PsAID) questionnaire, a 13-country EULAR initiative. Annals of the rheumatic diseases. 2014;73(6):1012-9.
6. Maruish, M. E. (Ed.). *User's manual for the SF-36v2 Health Survey* (3rd ed.). Lincoln, RI: QualityMetric Incorporated.

11 APPENDICES

11.1 APPENDIX 1: DEFINITIONS AND DATA DERIVATION DETAILS

- **Derivation of subgroups**
  - BMI = Weight (kg) / [Height (m)]^2, where:
    - Weight (kg) = 0.45359237 × Weight (lb).
    - Height (cm) = 2.54 × Height (in).
  - Waist circumference (cm)= 2.54 × Waist circumference (in).
  - Treatment line
    Patients will be classified into three groups based on treatment received for PSA at enrolment: Tofacitinib as monotherapy, Combination therapy with Methotrexate [MTX], and Combination with other csDMARD using the following steps:
    1. The start and end dates of Tofacitinib treatment are to be determined when the sequence number is 1.
    2. It is to be examined whether these dates fall within the treatment period of Methotrexate (MTX) or any other conventional synthetic Disease-Modifying Anti-Rheumatic Drugs (csDMARDs).
    3. If an overlap is found, the initial treatment is to be classified as a combination therapy, either with MTX or with other csDMARDs, as applicable.
    4. If no overlap is found, the initial treatment is to be classified as Tofacitinib monotherapy.

  - For both ESR and CRP, result will be considered Abnormal if t the test result is above the normal range.
    - For CRP the normal range is: 0.3 to 10 mg/L
    - For ESR the normal range is:

Age	Male	Female
0-50	<15 mm/h	<20 mm/h
51-85	<20 mm/h	<30 mm/h
>85	<30 mm/h	<42 mm/h

  - Unequivocal radiological erosion will be derived and presented as a categorical variable defined as follows: Present, if Sharp-Van der Heijde modified score for erosion >0; Absent, if Sharp-Van der Heijde modified score for erosion score = 0.
  - Arthritis at enrollment is identified when a patient report swelling and/or tenderness/pain in any joint. The number of arthritic joints in every joint category is determined by comparing the count of swollen joints to that of tender/painful joints and choosing the greatest value.
  - Size of the joint is determined as follows:
    - Large joints refer to glenohumeral, elbows, hips, knees and ankles joint.

  - Medium joints refer to temporomandibular, sternoclavicular, acromioclavicular, finger proximal interphalangeal, finger distal interphalangeal and tarsus/midfoot (feet) joints.
  - Small joints refer to metacarpal phalangeal joints (MCP) joints, proximal interphalangeal (PIP) joints, second through fifth metatarsal phalangeal joints (MTP) joints, thumb interphalangeal (IP) joints and wrists.
- Hand joints refer to MCP, finger proximal interphalangeal and finger distal interphalangeal joints.
- Baseline measurement is defined as the measurement taken at enrollment.
- Disease duration.
  Psoriatic Arthritis (PsA) disease duration will be calculated from the date of diagnosis to the enrollment visit date obtained from the electronic Case Report Form (eCRF) and presented in years.
  - PsA disease duration (years) = (enrollment visit date – date of first study date of diagnosis + 1)/365.25.
- Tofacitinib exposure
  - Time from enrollment to first dose (days) = first treatment administration date – enrollment date + 1.
  - Treatment duration (months) = (last treatment administration date – first treatment administration date + 1)/30.44.
- PSA Treatment will be classified as prior or current concomitant as follows:
  1. If the end date is prior to the enrolment date, then it is prior medication.
  2. If either the start or end date is after the enrolment date, or if it is reported as ongoing, then the medication is classified as current concomitant.
  3. If the start date is prior to the enrolment date, and the end date is after the enrolment date, the medication is classified as both prior and current concomitant medication.
  4. If the start date is prior to the enrolment date, and the end date is missing, the medication is classified as prior.
  5. If the start date or end date is missing and is not classified in any of the aforementioned categories, it is categorised as missing.

- **Psoriatic Arthritis Disease Activity Score (PASDAS)**

  The PASDAS (1) is a composite disease activity measure (range 0–10) for PsA. The PASDAS incorporates assessment of joints, dactylitis, enthesitis, physical function, Quality of Life (QoL), acute-phase response, and both patient and physician global ratings of disease. A PASDAS score between 0-10 is calculated using a weighted formula (0= no disease, 10= severe disease), as follows:

  PASDAS = (0.18 × √Physician global assessment Visual Analog Scale [VAS]) + (0.159 × √Patient global assessment VAS) – (0.253 × √SF-36 physical component score) + (0.101 × ln [Swollen joint count [SJC66] + 1]) + (0.048 × ln [Tender joint count [TJC68] + 1]) + (0.23 × ln [Leeds enthesitis count [LEI] + 1]) + (0.377 × ln[Leeds dactylitis count [LDI] + 1]) +((0.102 × ln [CRP (mg/L) + 1)] + 2) × 1.5)
  PASDAS VAS are based on a 100 mm.

- **The Leeds Enthesitis Index (LEI)**

  The LEI (2) aims to assess enthesitis in patients with PsA. Tenderness on examination is recorded as either present (1) or absent (0) for each of the 6 sites. The classification in each one of the sites is summed leading to the LEI overall score (range of 0–6), where higher count represents greater enthesitis burden.

- **Disease Activity in Psoriatic Arthritis (DAPSA)**

  The DAPSA (2, 3) is a composite disease activity measure and is calculated as follows:
  DAPSA = Swollen joint count (SJC66) + Tender joint count (TJC68) + Patient global assessment VAS + Patient pain assessment VAS + CRP (mg/dL)
  DAPSA VAS are based on a 10 cm scale.

- **Psoriatic Arthritis Impact of Disease 12 Questions (PsAID12)**

  The PsAID12 (5) is a questionnaire used for evaluating how much psoriatic arthritis impacts QoL and is comprised by the following domains: Pain, Fatigue, Skin problems, Work and/or leisure activities, Functional capacity, Discomfort, Sleep disturbance, Coping, Anxiety, fear and uncertainty, Embarrassment and/or shame, Social participation and Depression. Each one of the domains is based on a 0–10 numerical rating scale (NRS) and with a different weight.
  The range of the final PsAID12 value is 0–10 where higher numbers indicate worse status.
  PsAID12 = (PsAID12.Q1 [pain] NRS value ×3) + (PsAID12.Q2 [fatigue] NRS value ×2) + (PsAID12.Q3 [skin] NRS value ×2) + (PsAID12.Q4 [Work and/or leisure activities] NRS value ×2) + (PsAID12.Q5 [function] NRS value ×2) + (PsAID12.Q6 [discomfort] NRS value ×2) + (PsAID12.Q7 [sleep] NRS value ×2) + (PsAID12.Q8 [coping] NRS value ×1) + (PsAID12.Q9 [anxiety] NRS value ×1) + (PsAID12.Q10 [embarrassment] NRS value ×1) + (PsAID12.Q11 [social life] NRS value ×1) + (PsAID12.Q12 [depression] NRS value ×1)
  The total is divided by 20. Thus, the range of the final PsAID value is 0-10 where higher figures indicate worse status.
  If one of the 12 NRS values composing the PsAID is missing, the imputation is as follows: calculate the mean value of the 11 other (non-missing) NRS (range, 0-10) impute this value for the missing NRS. Then, calculate the PsAID as explained above.
  If 2 or more of the NRS are missing, the PsAID total score is considered missing (no imputation).

- **Spondylarthritis Research Consortium of Canada Enthesitis Index (SPARCC-EI)**

  The SPARCC-EI evaluates 16 enthesial sites: Greater trochanter (R/L), Quadriceps tendon insertion into the patella (R/L), Patellar ligament insertion into the patella and tibial tuberosity (R/L), Achilles tendon insertion (R/L), Plantar fascia insertion (R/L), Medial epicondyles (R/L), Lateral epicondyles (R/L) and Supraspinatus insertion (R/L) for the presence or absence of tenderness.

Tenderness at each site is quantified as: 0 = non-tender and 1 = tender. The maximum score of the SPARCC-EI is 16.

- **Psoriasis Area Severity Index (PASI)**

  The PASI (2) is an index used to express the severity of psoriasis, combining the severity (erythema, induration and desquamation) and percentage of affected area. Score range is 0 –72 and is calculated as follows.

  The score calculations will be split in 3 parts.
  For each body region (head, upper limbs, trunk and lower limbs) the plaque characteristics will be classified according to erythema, induration/thickness and scaling using the following lesion scores:
    - None = 0
    - Slight = 1
    - Moderate = 2
    - Severe = 3
    - Very Severe = 4

  A scores are calculated by adding together the three scores (erythema, induration/thickness and scaling) for each body region (head, upper limbs, trunk, and lower limbs) . B scores measure the area of involvement in each body region. They are graded from 0 to 6 depending on the estimated percentage of lesioned area as follows:
    - 0% = 0
    - 1% - 9% = 1
    - 10% - 29% = 2
    - 30% - 49% = 3
    - 50% - 69% = 4
    - 70% - 89% = 5
    - 90% - 100% = 6

  C score are obtained by multiplying A and scores for each body region. To obtain D scores, C scores are multiplied by the weight of the respective body region (0.1 for head and neck, 0.2 for upper limbs, 0.3 for trunk, and 0.4 for lower limbs). Finally, the PASI total score is calculated by summing the four D scores.

- **36-Item Short Form Health Survey questionnaire (SF-36)**

  The 36-Item Short Form Health Survey questionnaire (SF-36) is a questionnaire used for evaluating Health-Related Quality of Life (HRQoL) (6).
  The SF-36 measures a patient's functional health and well-being based on 36 questions about physical health (21 items) and mental health (14 items) concepts that are relevant across age, disease, and treatment.

  These questions are combined into groupings to form 2 summary measures (Physical Health and Mental Health) and 8 scales: the Physical Health Component Summary and its 4 scales (physical functioning, role-physical, bodily pain, and general health); and the Mental Component Summary and its 4 scales (vitality, social functioning, role-emotional, and mental health). The raw scores are recoded to a 100-point scale using a scoring key and summary measure scores are calculated. A higher score represents better HRQoL. Although item 2

pertains to overall general health, it is not included in any global summary measure. But it is used to understand a patient's perceived health in comparison to the previous year. Item response options range from 1 (Much better now than one year ago) to 5 (Much worse now than one year ago). Please refer to GenSight manual or Ware et.al. (Ware, Kosinski, Turner-Bowker, Gandek, & Maruish, 2007) for details.

The Physical Health Component Summary (PCS) comprises the following 4 subscales (or "domains"):

- The 10-item physical functioning subscale assesses physical activities in a typical day and the extent at which they are limited. Item response options are 1 (yes, limited a lot), 2 (yes, limited a little), or 3 (no, not limited at all). Low scores represent significant limitations with activities and high scores indicate little or no limitations.
- The 4-item role participation with physical health problems/role-physical subscale assesses how often the patient's physical health has impacted the performance of tasks and daily activities in the past 4 weeks. Item response options range from 1 (all of the time) to 5 (none of the time). Low scores indicate more problems with activities, and high scores indicate no problems.
- The 2-item bodily pain subscale assesses intensity of pain and interference of pain with normal work activity in the past 4 weeks. Response options for pain severity range from 1 (none) to 5 (severe). Response options for pain interference range from 1 (not at all) to 5 (extremely). High scores indicate high pain that impacts activities, and lower scores indicate no pain and no impact on normal activity. Therefore, the scores for both items are to be properly recoded to calculate the subscale score for which a higher score indicates better status.
- The 5-item general health subscale includes 4 items that address the patient's expectations of health, and one item that addresses self-rating of health. Response options for expectation of health range from 1 (definitely true) to 5 (definitely false). Response options for rating of health range from 1 (excellent) to 5 (poor). High scores indicate negative self-assessment of health and low scores indicate positive self-assessment of health. Therefore, the scores for both items are to be properly recoded to calculate the subscale score for which a higher score indicates better status.

The Mental Component Summary (MCS) comprises the following 4 subscales (or "domains"):

- The 4-item vitality subscale assesses energy levels in the past 4 weeks. Item response options range from 1 (all of the time) to 5 (none of the time). Low scores indicate feeling tired and worn out, and higher scores indicate feeling full of energy most of the time.
- The 2-item social functioning subscale assesses quantity and quality of social activities in the past 4 weeks. Item response options for extent of

interference on quantity of social activities range from 1 (all of the time) to 5 (none of the time). Item response options for extent of interference on the quality of social activities range from 1 (not at all) to 5 (extremely). High scores indicate extreme physical and emotional impact, and low scores indicate no impact. Therefore, the scores for both items are to be properly recoded to calculate the subscale score for which a higher score indicates better status.

- The 3-item role participation with emotional health problems/role-emotional subscale assesses mental health during the past 4 weeks. Item response options range from 1 (all of the time) to 5 (none of the time). Low scores indicate problems with work or activities because of emotional issues, and high scores indicate no limits on work or activities.
- The 5-item mental health subscale assesses anxiety, depression, loss of behavioral/emotional control, and psychological well-being during the past 4 weeks. Item response options range from 1 (all of the time) to 5 (none of the time). Low scores indicate feelings of nervousness and depression, and high scores indicate feelings of peace and happiness.

Please refer to Table 2 and Table 3 for recoding and scoring algorithms, respectively. Within each PCS or MCS subscale, if more than half of the items are missing (e.g. not answered, 2 or multiple options selected, and/or out-of-range value entered) and not imputable, the subscale score will be set to missing. Otherwise, the subscale scores will be calculated as follows (after all item scores are properly recoded / reversed if necessary):

1. Each missing item will be imputed with average of the other non-missing items.
2. The sum of all item scores will be set as the raw subscale score.
3. The raw subscale score obtained from step 2 will be transformed using the formula: Transformed subscale score = 100 × (raw subscale score – lowest possible subscale score) / possible subscale score range.

For example, for physical functioning, the lowest possible subscale score is 10 (items) × 1 = 10, while the highest possible subscale score is 10 (items) × 3 = 30, so the range is 30 - 10 = 20. If the raw physical functioning score is 21, the transformed physical functioning score will be 100 × (21-10)/20 = 55.
The transformed subscale score can be converted into Z-score and T-score. Furthermore, each transformed subscale score will be assigned a weighting factor to calculate PCS and MCS Summary composite scores. Please note, however, that the 4 PCS subscales also contribute to MCS composite score, and vice versa. Details are illustrated in Table 4.

Table 1. SF-36 (v2) Scoring Key: Item Value Recoding

Item Number	Original Response Number	Recoded Value
1	1	5
	2	4.4
	3	3.4
	4	2

Item Number	Original Response Number	Recoded Value
	5	1
7	1	6
	2	5.4
	3	4.2
	4	3.1
	5	2.2
	6	1
8	1	6 (or 5*)
	2	4 (or 4.75**)
	3	3 (or 3.5**)
	4	2 (or 2.25**)
	5	1
3a-3j	1	1
	2	2
	3	3
4a-4d, 5a-5c, 9b, 9c, 9f, 9g, 9i, 10, 11a, 11c	1	1
	2	2
	3	3
	4	4
	5	5
6, 9a, 9d, 9e, 9h, 11b, 11d	1	5
	2	4
	3	3
	4	2
	5	1

* If Item 7 ≥ 2.
** If and only if Item 7 is missing.

Table 2. SF-36 (v2): Item(s) to Generate Subscale Scores

Subscale	Item No.	Number of Items to be Summed	Raw Subscale Sum		
			Minimum	Maximum	Range
Physical Functioning	3a-3j	10	10	30	20
Role-Physical	4a-4d	4	4	20	16
Bodily Pain	7, 8	2	2	12	10
General Health	1, 11a-11d	5	5	25	20
Vitality	9a, 9e, 9g, 9i	4	4	20	16
Social Functioning	6, 10	2	2	10	8

Role-Emotional	5a-5c	3	3	15	12
Mental Health	9b-9d, 9f, 9h	5	5	25	20
Self-Evaluated Transition	2	N/A	N/A	N/A	N/A

Transformed Subscale Score = 100 × (Raw Subscale Sum – Minimum)/Range.
Health Transition should always be considered ordinal and no transformation is necessary.

Table 3. SF-36 (v2): Parameters and Weighting Factors to Generate PCS and MCS

				Weighting Factor	
i	Transformed Subscale Score (a_i)	μ_i	s_i	PCS (w_{1i})	MCS (w_{2i})
1	Physical Functioning	83.29094	23.75883	0.42042	-0.22999
2	Role-Physical	82.50964	25.52028	0.35119	-0.12329
3	Bodily Pain	71.32527	23.66224	0.31754	-0.09731
4	General Health	70.84570	20.97821	0.24954	-0.01571
5	Vitality	58.31411	20.01923	0.02877	0.23534
6	Social Functioning	84.30250	22.91921	-0.00753	0.26876
7	Role-Emotional	87.39733	21.43778	-0.19206	0.43407
8	Mental Health	74.98685	17.75604	-0.22069	0.48581

Z-score $z_i = (a_i - \mu_i)/ s_i$, i=1, 2, …, 8 PCS = $50 + 10\times\{S_{i=1 \text{ to } 8} (z_i \times w_{1i})\}$, i=1, 2, …, 8
T-score $t_i = 50 + 10\times z_i$, i=1, 2, …, 8 MCS = $50 + 10\times\{S_{i=1 \text{ to } 8} (z_i \times w_{2i})\}$, i=1, 2, …, 8

- **Health Assessment Questionnaire – Disability Index (HAQ-DI)**

  The HAQ-DI (4) is a health questionnaire assessing the degree of difficulty a patient has experienced during the past week in 8 domains of daily living activities. From the HAQ-DI the following components are displayed: dressing and grooming (2 questions), arising (2 questions), eating (3 questions), walking (2 questions), hygiene (3 questions), reach (2 questions), grip (3 questions), activities (3 questions), and pain (from VAS).
  The score for each component of the HAQ is based on the following validated scoring system: without any difficulty (0), with some difficulty (1), with much difficulty (2), unable to do (3).

  The highest score for any component question determines the score for that category.

  If a component question is left blank or the response is too ambiguous to assign a score (i.e. not applicable), then the score for that category is determined by the remaining completed question(s). If all component questions are blank, then the category is left blank.

For each component there are also some devices that can impact component score derivation. Full details of each component scoring are reported in the Table 1.

Table 4. HAQ components scoring details

Component	Questions	Aid/Device	Component Score Derivation
Dressing & Grooming	2	• Devices used for dressing (button hook, zipper pull, • Long-handled shoehorn, etc.)	Set score to the maximum of each question. If aid/device or help of another person is indicated set score =2 unless any of the 4 questions have a value of 3, in this case the score = 3.
Arising	2	• Special or built up chair	
Eating	3	• Built up pencils or special utensils	
Walking	2	• Cane • Walker • Crutches • Wheelchair	
Hygiene	3	• Raised toilet seat • Bathtub seat • Bathtub bar • Long-handled appliances in bathroom	
Reach	2	• Long-handled appliances for reach	
Grip	3	• Jar opener (for jars previously opened)	
Activities	3	• Not applicable	

The disability index will be computed by adding the scores for each of the components and dividing by the number of components with an available score. A disability index cannot be computed if the patient does not have scores for at least six (6) categories.

11.2 APPENDIX 2: DEFINITION AND USE OF VISIT WINDOWS IN REPORTING

Data will be summarized in tables by visit when applicable. The following visit label and visit windows will be applied for the analysis of effectiveness and PRO endpoints:

Visit	Visit Label	Target Day	Visit Window
Visit 1	Baseline	1	Day 1 or before
Visit 2	Month 3	91	Day 2 to Day 137
Visit 3	Month 6	183	Day 138 to Day 274
Visit 4	Month 12	365	Day 275 onwards

To meet the requirements of sample collection, the Baseline visit window extends from day 8 or before and the visit window for Month 3 visit extends from day 9 to day 137.
In case of multiple observations falling within a given window, the observations selected for analysis will be identified as follows:

1. The observation closest to the target day will be used.
2. If the observations are at equal distance from the target day in absolute value, the one with a correct nominal visit label will be used.
3. If neither (1) nor (2) can be used to identify the observation windowing, then the latest observation within the analysis window will be used.