Statistical Analysis Plan

For the

3TR PARTNER-RA Study

Version Number	Version 1.0 dated 13th February 2025
Approval Date	4th March 2025
Effective Date	4th March 2025
Author(s)	Rajnikant Mehta, Trial Statistician Rhian Gabe, Lead Trial Statistician
Review by	Joseph Dib, Trial Manager Liliane Fossati-Jimack, Senior Scientific Officer Costantino Pitzalis, Chief Investigator
Approved by	Rhian Gabe, Lead Statistician

Centre for Experimental Medicine and Rheumatology

William Harvey Research Institute

Barts and the London School of Medicine and Dentistry

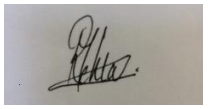
SAP revision history

Version	Section number changed	Description of change	Date changed	Date approved	Author
0.1	N/A	N/A	N/A	N/A	Rajnikant Mehta
1.0	Proof read reviewed	Final copy non draft	13.02.2025		Rajnikant Mehta, Joseph Dib

Full Trial title: The **3TR** Molecular **PA**thobiology and P**R**ecision **T**herapy i**N** **E**a**R**ly **R**heumatiod **A**rthritis (3TR PARTNER-RA) Study

Short title/Acronym:	3TR PARTNER-RA
Chief Investigator	Prof Costantino Pitzalis (CP)
Trial Manager	Jospeh Dib (DP)
Lead Statistician	Prof Rhian Gabe (RG)
Trial Statistician	Rajnikant Mehta (RM)

Signatures

Rajnikant Mehta ***15th January 2025:*** 

Trial Statistician Date and Signature

Rhian Gabe ***15th* January 2025**

Lead Statistician Date and Signature

Costantino Pitzalis

Digitally signed by Costantino Pitzalis
Date: 2025.03.04 09:53:19 Z

Chief Investigator Date and Signature

Table of Contents

Abbreviations and Definitions

ACR	American college of rheumatology
ANCOVA	Analysis of covariance
AUC	Area under the curve
CDAI	Clinical disease activity index
CRP	C-reactive protein
CSR	Clinical study report
CTU	Clinical Trials Unit
DAS	Disease activity score
csDMARD	Conventional synthetic Disease Modifying Anti-Rheumatic Drugs
EMR	Experimental Medicine and Rheumatology
EQ-5D-5L	EuroQol- 5 Dimension Level 5
EULAR	European League Against Rheumatism
ESR	Erythrocyte sedimentation rate
ESS	Epworth Sleepiness Scale
FACIT	Functional Assessment of Chronic Illness Therapy
FDC	Follicular dendritic cell
HAQ	Health Assessment Questionnaire
IQR	Interquartile range
ISRCTN	International Standard Randomised Controlled Trial Number
ITT	Intention to Treat
MRC	Medical Research Council
MTR	Major treatment response
MTX	Methotrexate
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NRAS	National Rheumatoid Arthritis Society
PI	Principal Investigator
PPI	Patient and Public Involvement

PP	Per protocol
QMUL	Queen Mary University of London
RA	Rheumatoid arthritis
SAF	Safety analysis set
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SF-36	36-Item Short Form Survey (SF-36)
TNFi	Tumour necrosis factor inhibitor
US	Ultrasound
VAS	Visual Analogue Scale

1. Statistical analysis plan (SAP) authorship

The Trial Statistician is responsible for writing the Statistical Analysis Plan (SAP), writing the computer code for implementing the SAP at the point of analysis.

The Trial Manager and the Lead Trial Statistician will review the SAP. The Statisticians and the Chief Investigator will sign off the final SAP.

The SAP has been finalised prior to unblinded analysis. The SAP has been drafted by the Trial Statistician and approved by the Lead Statistician. The Trial Statistician will complete the randomisation checks, TSC reports (open reports) and complete the final trial analysis. If any changes are requested to the SAP during the trial, this should be confirmed and agreed in writing by the Lead Statistician and Independent Statistician. Following trial analysis, the primary analysis will be checked by the Lead Statistician or an independent Statistician (independent from the study team and has not been involved in any 3TR PARTNER-RA data analysis previously).

Blinding within the trial:

The participant nor any of the investigators or site staff who are involved in the treatment or clinical evaluation and monitoring of the participants will be aware of the study intervention received. The Chief Investigator and recruiting study site teams will be blinded to the arm that the patient has been allocated to (intervention or control arm).

This document is the SAP for the 3TR PARTNER-RA trial and should be read in conjunction with the current trial protocol. This SAP details the proposed analyses and presentation of the data for the main paper(s) reporting the results for the 3TR PARTNER-RA trial.

The results reported in these papers will follow the strategy set out here. Subsequent analyses of a more exploratory nature will not be bound by this

strategy, though they are expected to follow the broad principles laid down here. The principles are not intended to curtail exploratory analysis (e.g. to decide cut-points for categorisation of continuous variables), nor to prohibit accepted practices (e.g. transformation of data prior to analysis), but they are intended to establish rules that will be followed, as closely as possible, when analysing and reporting data.

Any deviations from this SAP will be described and justified in the final report or publication of the trial (Appendix A). The analysis will be carried out by an appropriately qualified statistician, who should ensure the integrity of the data during their data cleaning processes.

2. Introduction

The 3TR PARTNER-RA trial is a double blinded, phase IV, randomised controlled, multi-centre, controlled Clinical Trial of an Investigational Medicinal Product (CTIMP) in a population of patients newly diagnosed with RA (symptoms <12 months) and fulfilling the 2010 ACR/EULAR classification criteria for RA who have an inadequate response to conventional Disease Modifying Anti-Rheumatic Drugs (DMARDs) and are eligible for anti-TNF therapy according to EULAR recommendations: treatment for ≥3months with ≥ 1 csDMARDs[1]. Patients recruited to this study will undergo a synovial biopsy at baseline and randomised (1:1) to drug (abatacept) and methotrexate or placebo and methotrexate arm. Patients will be followed up to 16 weeks at which point all patients will be offered a voluntary second synovial biopsy.

In recent years, thanks to the development of minimally invasive ultrasound-guided biopsy[2] synovial tissue analysis has been proposed as a tool for patient stratification with potential use in clinical practice. In patients with early RA, recent publications indicated that histological[3] and molecular[4] signatures in the synovial tissue are associated with disease outcomes and can stratify response to treatment and predict disease progression, including use of biologics[5]. Recently, the first biopsy-driven Randomised Clinical Trials in RA has shown that the lack of B cell lineage signatures in synovia is associated with lack of response to B cell depleting agent (rituximab) as compared to an alternative medication targeting IL6 receptor (Tocilizumab)[6].

Patients will be randomised to either an intervention arm or control arm after having a synovial biopsy. If randomised to the intervention arm (treatment by biomarker), the patient's biopsy tissue will be analysed within 2 weeks. In the absence of a target biomarker the patient will be randomised 1:1 to either abatacept or placebo. If the patient is randomised to the control arm, they will be randomised again 1:1 to either Abatacept + MTX or placebo. Patients will continue trial treatment

until 16 weeks, where the treatment response will be assessed as the primary endpoint. There will be a post-treatment visit/call scheduled 30+ days after week 16 (Visit 7). Patients will attend 5 visits (Protocol: Study Schedule, Page 30).

The end of the study will be triggered up to a maximum of 4 months after the last patient completes their final study visit (Last Patient Last Visit-LPLV) at the post-treatment visit/call assessment 30+ days following completion of trial treatment (Protocol Section 13.1).

All joint assessments (Protocol Section 9.8) will be performed by a member of the local trial team who will be blinded to treatment allocation. All participating site staff will be blinded to the arm of the study the patient has been allocated to (Protocol Section 9.8).

The specific aim of the sub-study is to discover if a biopsy at presentation prior to any RA treatment (collected as part of the 3TR Early RA study) can be used to determine treatment response even after the patient has received cDMARDs.

2.1 Study objectives

Primary objective: The primary objective of this trial is to determine whether synovial molecular profiles (=drug target signatures) can inform treatment response to abatacept in early RA..

To this aim, we will compare the change in Clinical Disease Activity Index (CDAI) score at 16 weeks between the biomarker positive and the biomarker negative patients within the abatacept group (i.e. Group 1 vs Group 3) Fig.1 Study Scheme diagram. The primary trial hypothesis is that biomarker positive patients treated according to their highest expressed biomarker will have lower CDAI response at 16 weeks compared to biomarker negative patients.

This primary analysis has been chosen as we want to establish a difference in response rates in the treatment in the biomarker arm first, and the trial has been powered for this.

Secondary objectives:

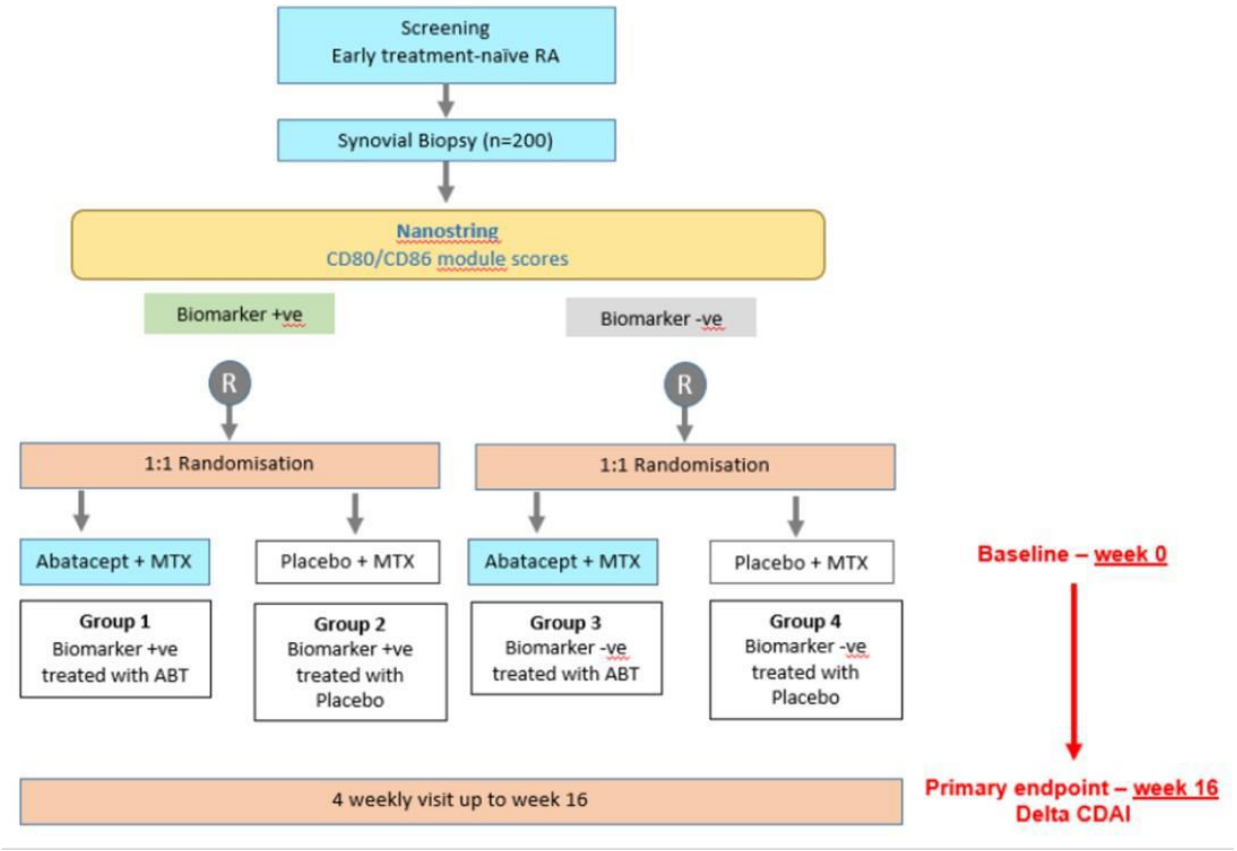
The following analysis will be done:

I. Compare patients treated according to their biomarker in the intervention arm against the control arm as a whole (i.e. Group 1 vs 3+4), to determine the enrichment in response in the treatment allocation arm vs the standard of care response rate.

II. Assess the efficacy of treatment allocation according to biomarker compared to random allocation, we will compare the biomarker positive patients in the intervention arm vs the biomarker positive patients in the control arm (i.e. Group 1 vs Group 4)

III. Assess the efficacy of the strategy as a whole against the current clinical practice by comparing the control vs intervention arm (i.e. Groups 1 + 2 vs Group 3+4)

2.2 Study scheme diagram



The primary outcome will be assessed at 16 weeks from baseline.

2.3 Sample size determination

The primary aim of the study is to demonstrate that patients in the "Drug target biomarker" group (Group 1) will have a higher response rate than the "No drug target biomarker" groups (Group 2 and 3)- see section 2.2 for trial diagram.

Using change in CDAI score from baseline (delta CDAI) as the primary outcome, a difference in delta CDAI between the two comparison groups of 10, with a Standard Deviation of 14.34. These are the values that were observed when analysing the difference in delta CDAI between biomarker positive and biomarker negative patients treated with MTX in the PEAC study. Although PEAC patients were not treated with Abatacept, this comparison is expected to be even more in favour of the biomarker positive group using Abatacept together with MTX. Thus, basing the power calculation on the PEAC observed difference is a conservative assumption.

With these assumptions, 90 patients would be needed to detect a significant difference of 10 in delta CDAI with 90% power. Since this would only be for Group 1 and Group 3 (45 patients each), a total of 180 patients would be sufficient for 90% power in the whole trial (Figure 1 Section 2.6). If we assume a 5% dropout rate and also account for an expected 5% of patients that have insufficient RNA for the Nano-string analysis (based on results from the STRAP trial), 200 patients in the whole trial would be sufficient to detect a significant difference (Group 1 vs Group 3) of 10 in delta CDAI with 90% power.

All power calculations were performed using the "twomeans" package in STATA, version SE 18.

2.4 Randomisation

Randomisation will take place when all the screening procedures and the biopsy visit are completed, and the patient is eligible for enrolment in the study. The randomisation in the study will be applied using simple randomisation, with equal allocation ratio (1:1). All patients with sufficient RNA for analysis will be randomised using simple randomisation (1:1 allocation ratio) into the Control or the Treatment arm according to biomarker arm.

Patients allocated to the control arm of the study, will be subsequently randomised again using simple randomisation with equal treatment allocation ratio (1:1) to Abatacept and Methotrexate therapy or placebo and Methotrexate therapy.

Once the biopsy sample has been received by the central laboratory, patients will be stratified into 2 groups, CD80/CD86 positive (biomarker+) or CD80/CD86 negative (biomarker-) and randomised 1:1 to receive either Abatacept and Methotrexate therapy or placebo and Methotrexate therapy.

3. General Analysis Definitions

3.1 Study timelines and visit definitions

3.1.1 Study timelines

Data cleaning will be carried out periodically throughout the trial duration and final cleaning will take place up to 4 months after Last Patient Last Visit (LPLV). The main analyses will be carried out following database lock, in time for the clinical study report which will be submitted to the REC. The trial results will be uploaded to the EudraCT website.

3.1.2 Visits

Patients who consent are screened and undergo a synovial biopsy. Randomisation will take place when all the screening procedures are completed, and the biopsy is performed, and the patient is eligible for enrolment in the study (Protocol section 6).

A synovial biopsy will be carried out for each patient before the baseline visit (Protocol section 2.6). This initial synovial biopsy is mandatory as part of the patient stratification process; however, the subsequent week 16 synovial biopsy will remain optional.

Patients receive Abatacept + MTX or Placebo + MTX at Baseline (week 0) visit 3 depending upon randomisation outcome. All baseline assessments are performed prior to commencement of therapy.

After baseline, all patients will be monitored on a 4 weekly basis (± 7 days).

3.2 Study populations

3.2.1 ITT population

ITT population includes all the patients who were enrolled in the trial and were randomised/assigned treatment according to biomarker. All patients will be analysed according to the randomisation/treatment allocation including any withdrawals after randomisation/treatment allocation.

There are two types of withdrawal defined; full withdrawal from the study (the patient does not attend any further visits) or withdrawal from trial treatment (treatment cessation). The data collection and follow up requirements are specified in Sections 15 of the trial protocol.

ITT analysis will be used for all analyses except for the safety analysis in Section 14.9.

3.2.2 Other populations

Definition/Description of Other population is in the Protocol Section 14.6

3.2.3 Safety analysis set (SAF)

Safety analysis set consists of patients who received at least one dose of the trial medication. This population is identical to the ITT population except that patients will be analysed according to their actual treatment in case this differs from the scheduled treatment. This population set will be used to report the safety data.

3.3 Study Endpoints

3.3.1 Primary endpoint

The primary endpoint is the change in CDAI between baseline and at 16 weeks. This endpoint is a continuous outcome.

A responder is defined as a patient who achieved an improvement of 10 units in CDAI: 10 unit reduction is regarded and an improvement.

3.3.2 Secondary endpoints

1) Percentage of patients with DAS28<3.2 (LDA) at 16 weeks.
2) Percentage of patients deemed responders using American College of Rheumatology 50 (ACR50) measure at 16 weeks.
3) Percentage of patients with CDAI remission at 16 weeks.
4) Change in HAQ-DI at 16 weeks from baseline.
5) Change in SF-36 at 16 weeks from baseline

3.3.3 Exploratory endpoints

1) Percentage of patients deemed responders using American College of Rheumatology 20/70 (ACR20/70) measure at 16 weeks.
2) Percentage of patients with ACR/EULAR Boolean remission at 16 weeks.
3) Mean % change in CDAI score and DAS28 at 16 weeks.

4) Change in FACIT, ESS, and EQ-5D score at 16 weeks from baseline

5) The association between synovial histology and ultrasound measures of inflammation, drug response rates, disease outcome and disability.

6) SAEs from 0 to 16 weeks for all patients

7) Change in 12-max summary measure of US 2D synovial thickness (ST) grey scale and power Doppler (PD) signal from baseline at 16 weeks.

8) Changes from baseline to week 16 in the total histopathological synovitis score.

4. Descriptive statistics

All the following descriptive statistics of patients in the study (will be presented separately and then combined.

Demographics and patient characteristics are listed in the table below. For further information, please see Protocol Section 7.5

Table 1: List of demographics and characteristics.

Variable type	Variables
Demographics	Age, Gender, Ethnicity, Medical History
Diagnostics	2010 ACR/EULAR RA classification criteria (joint involvement, serology, acute phase reactant, symptom duration), rheumatoid factor, CCP, disease duration
Disease activities	ACR/EULAR core set, DAS28, CDAI, VAS
Physical function	HAQ,
Fatigue score	FACIT
Health status	SF-36, EQ-5D-5L
Sleepiness	ESS
Vital signs	Weight, Height, BMI, blood pressure (systolic, diastolic), pulse, temperature
Cardiovascular risk factors	Alcohol status, smoking status and smoking years, lipid lowering agent, ischemic heart disease, diabetes, hypertension, family history
Routine blood tests	ESR, CRP, Hb, Haematocrit, WBC, Platelets, Neutrophils, Lymphocytes, ALT, AST, Creatinine, Urea, Sodium, Potassium, eGFR, Red blood cell count, Mean cell haemoglobin, Mean cell volume
Image assessments	Radiograph scores (Sharp van der Hyde scores, including erosion scores, joint space narrowing and total) Ultrasound scores (Power doppler and Synovial thickening)

4.1 Demographics and baseline characteristics

The summary of continuous variables will include the number of subjects, the mean and standard deviation, the median, the minimum and the maximum as well as the numbers with missing values. For categorical variables, the frequencies and percentages will be presented, where the percentages will be based on the complete cases (without missing data). A summary will be presented overall, by biomarker (positive/negative) and by treatment. At baseline, the summary is applied on each of the variables listed in Table 1.

4.2 Characteristics collected at 16 weeks

The same summary as for the baseline will be carried out at 16 weeks, by biomarker and by treatment.

4.3 Prior and concomitant therapies

Prior and concomitant therapies are defined in the Protocol Section 11.15.

4.4 Baseline biomarker

Baseline biomarker will be summarised using frequency and percentage and displayed by treatment.

4.5 Study drug

For each study drug exposure, number and percentage of patients will be reported by visit and dose received.

Adherence to the planned dose regimen of study medication will be summarised (dose modification, reduction, delay) by visit and treatment.

4.6 Protocol violations

All major protocol violations will be summarised by type of violation, treatments groups and pathotype.

4.7 General considerations

For each treatment group and biomarker, the following will be presented:

• Number and percentage of subjects screened, not randomised/assigned treatment, randomised/assigned treatment, randomised/assigned treatment and not started treatment, discontinued, and completed.

• Number and percentage of subjects randomised at each site.

• Number and percentage of subjects in each analysis set (ITT) overall and by visit.

• Number and percentage of subjects excluded from each analysis set by reason for exclusion.

• Number and percentage of withdrawals (withdraw consent to any further participation in the trial) and treatment cessation, by reason for discontinuation and visit.

5. Interim analysis and timing for analysis

No interim analysis is planned for the current trial. However, the TSC will review the accruing trial data every year during the recruitment phase of the study and assess whether there are any safety issues that should be brought to participants' attention or any reasons for the trial not to continue. TSC reports will consist of an open report where the accrual data will be described without any information on the arm patients have been allocated to, and a closed report where any safety issues by treatment arm will be presented. No stopping rule will be used. See section 9 for further detail. The trial data will be analysed following database lock.

6. Efficacy analysis

The listed measures in Table 2 are to be used in the analyses. The analyses will be carried out in STATA 18 or higher.

In hypothesis testing, the null hypothesis is rejected at (two-sided) significance level 0.05 unless otherwise specified.

6.1 Main analyses

6.1.1 Primary endpoint

The primary endpoint of the study will be a continuous outcome of CDAI measure at 16 weeks (visit 7).

The primary analysis will assess the difference in the CDAI comparing the difference in response between the abatacept and placebo groups with the biomarker positive group against the difference in response between the abatacept and placebo groups with the biomarker negative group.

This analysis will be on an ITT basis according to original randomisation allocation.

The differences between the groups on the primary outcome will be tested with a mixed-effects model with time as a continuous variable. The CDAI is measured at five time points: (1) Baseline, (2) 4 Weeks, (3) 8 Weeks, (4) 12 Weeks, (5) 16 Weeks. The baseline measurement for the primary outcome is the CDAI score after randomisation. This measurement will be included in the analysis as a covariate.

The mixed-effects model results will be presented as P values and 95% CI. In the mixed-effects model we will use both an 'unstructured' and a '1. order autoregressive' covariance matrix and choose the matrix resulting in the lowest Bayesian information criterion. To assess, if the underlying assumptions behind the mixed-effects model analysis are fulfilled, we will investigate normal quantile plots of residuals, standardized residuals, and random effects. If the underlying assumptions behind the mixed-effects model analysis are clearly violated then we will use a generalized estimation equation for the analysis.

The planned primary analysis is a mixed-effects model adjusted for centre and patient comparing the outcome of CDAI between treatment groups in the biomarker arm (Group 1) and non-biomarker arms (Group 3).

A p- value of less than 0.05 will be taken to indicate statistical significance.

6.1.2 Secondary analyses

The following secondary endpoints will be reported descriptively within each treatment group:

1) Percentage of patients with DAS28(ESR)<3.2 (LDA) at 16 weeks
2) Percentage of patients with CDAI ≤10 (LDA) at 16 weeks

3) Percentage of patients with CDAI remission at 16 weeks

4) Change in HAQ-DI at 16 weeks from baseline

5) Change in SF-36 at 16 weeks from baseline

Comparisons by groups will be guided by those performed for the primary endpoint in accordance with the following testing outlined above for the primary analysis:

1 vs **3**

1 + **2** vs **3+4**

All comparisons involving continuous secondary outcomes will be carried using mixed-effects models comparing treatment groups at 16 weeks, adjusted for baseline score, patient and centre. The associated absolute change difference and 95% confidence interval, as well as the mean change in each group will be reported.

6.2 Subgroup analyses

Subgroups analyses will be performed in individual treatment groups, both overall and by arm, will be presented descriptively.

6.3 Analysis of exploratory endpoints

The robustness of the results on the primary and secondary endpoints will be assessed using the exploratory outcomes as presented in the table below.

Table 2: List of exploratory outcomes that are used in the analyses.

Categorical disease outcomes		**Statistical analysis**
Percentage of patients with ACR 20 and 70 response rates at 16 weeks		Logistic regression with the treatment, and study as factors will be used. Exact logistic regression will be considered if the sample size is too small for a regular logistic regression and/or if some of the cells have no observations. Separate analysis will be performed for biomarker/no biomarker using the ITT population.
Continuous disease outcomes		**Statistical analysis**
Mean % percentage change in CDAI score and DAS28ESR at 16 weeks from baseline.	Change in 12-max summary measure of US 2D Synovial thickness (ST) and power Doppler (PD) signal at 16 weeks.	Mixed-effects modelling will be performed with treatment, time and patient as factors, and baseline score as covariate. If the assumptions of the Mixed-effects modelling are not met, the data will be analysed GEE methods.
Change in FACIT and ESS score at 12 weeks from baseline.	Changes from baseline to week 12 in the total histopathological synovitis score.	
SAEs analysis		**Statistical analysis**
SAEs from 0 to 16 weeks (+ 30 days –see section 7.3)		Frequency and percentage of each outcome will be reported by treatment arm and biomarker using the safety analysis set (SAF).
Association/correlation analysis		**Statistical analysis**
Association between disease outcome (binary variables) and synovial histology (binary variables)	Association between disease outcome (binary variables), treatment group and synovial histology (binary variables)	Pearson or Spearman (rank) correlation coefficients will be used to assess the correlation between two continuous measures depending on normality. T-test or Mann-Whitney test will be used to assess the association between a binary grouping and a continuous outcome depending on normality.
Association between disease outcome (binary variables) and synovial histology (continuous variables)	Association between disease outcome (binary variables), treatment group and synovial histology (continuous variables)	

Association between disease outcome (binary variables) and Ultrasound measure of inflammation	Association between disease outcome (binary variables), treatment group and ultrasound measure of inflammation	Two-way analysis of variance or Friedman's two- way analysis of variance will be used to assess the association between two categorical (including binary) measures and a continuous outcome depending on normality. Log linear model will be used to assess the association between three binary outcomes. Chi-square test/Fischer test will be used to assess the association between two binary outcomes.
Association between disease outcome (binary variables) and Disability	Association between disease outcome (binary variables), treatment group and disability	
Association between disease outcome (continuous variables) and synovial histology (binary variables)	Association between disease outcome (continuous variables), treatment group and synovial histology (binary variables)	
Association between disease outcome (continuous variables) and synovial histology (continuous variables)	Association between disease outcome (continuous variables), treatment group and synovial histology (continuous variables)	
Association between disease outcome (continuous variables) and Ultrasound measure of inflammation	Association between disease outcome (continuous variables), treatment group and Ultrasound measure of inflammation	
Association between disease outcome (continuous variables) and Disability	Association between disease outcome (continuous variables), treatment group and disability	

6.4 Generated variables

The generated variables are ACR20, ACR50, ACR70, CDAI, DAS28, HAQ, SF-36, FACIT-Fatigue, ESS and EQ-5D-5L scores.

Instrument	Transformation
ACR Improvement[7]	
20%	Refer to Protocol Section 9.8.1
50%	
70%	
CDAI[8]	Refer to Protocol Section 9.8.2
DAS28[9]	Refer to Protocol Section 9.8.2
HAQ[10]	Refer to Protocol Section 9.8.3
SF-36[11]	Refer to Protocol Section 9.8.4
FACIT-Fatigue[12]	Refer to Protocol Section 9.8.5
ESS[13]	Refer to Protocol Section 9.8.6
EQ-5D-5L scores[14]	Refer to Protocol Section 9.8.7

6.5 Assumptions for analysis

The independence of observations (patients) is assumed within centres, as well as between centres However, the intraclass correlation (ICC) between centres will be tested, if it is significant, a mixed-effects models regression will be used with random effects for centres (and fixed treatment effects).

6.6 Methods for handling dropouts, missing data and outliers

6.6.1 Handling of dropouts

The number of dropouts before 16 weeks will be monitored by treatment group. This will be reviewed at TSC meetings. If one treatment causes considerably higher dropouts, further investigation will be carried out to seek the explanations.

6.6.2 Handling of missing data

The missing values in primary outcome will be considered separately from other variables.

It is rare that the primary outcome is missing, but if so, the CDAI response at another visit will be used as the primary outcome. If a patient stops treatment before reaching visit 7 (when the primary endpoint is obtained) but CDAI response is available at visit 6, the CDAI at visit 6 will be used for the primary endpoint. When more than 5% of primary endpoint data is missing under other circumstances, Missing data for essential outcome variables and their subcomponents will be handled by multiple imputation by chained equation (MICE), incorporating longitudinal datapoints. Imputation will be performed five times and the mean (or median, if a particular variable is not Gaussian) result of imputation will be used to replace missing values. Imputed values will be checked for excessive variability across the five imputation runs assuming data are missing at random (MAR). Multiple imputations will be implemented using STATA (version 18 or higher). The performance of the imputations will be examined through the convergence and marginal distributions. Five imputations will be produced and the pooled results of the 5 analyses will be presented.

For questionnaires (SF-36, HAQ, EQ-5D-5L, ESS and FACIT fatigue), the database will produce the final scores. The missing component score values will be imputed using the same technique as describe above.

6.6.3 Handling of outliers

All values will be included in the analyses unless different decisions are made when reviewing the data.

Subject data will not be excluded from the analysis if the subject fails to comply with the visit schedule described in the protocol.

6.7 Statistical analysis issues

6.7.1 Multiple comparisons

The primary and secondary analyses will be carried at 5% significance level. Since one primary endpoint will be analysed, no correction for multiplicity will be applied. Many of the secondary outcome measures are expected to be highly correlated with the primary outcome measure, their analyses will be considered as confirmatory of the primary endpoint. Sub- group analyses are regarded as descriptive and will not be accompanied by formal hypothesis testing.

6.7.2 Multi-centre studies

The key outcome (CDAI) will be summarised using mean and standard deviation by centre.

If there is significant evidence that the centres are different from each other, the variation of centres will be adjusted in the models.

The statistical analyses will be performed using STATA (version 18 or higher). All applied tests will be two sided and p-values of 0.05 will be accepted as statistically significant.

6.7.3 Post-Hoc analysis

Post-hoc analyses may be performed to investigate, where possible, performance of alternative module cut points to inform future research.

7. Safety analysis

Safety analysis will be based on safety population (SAF). Protocol V2.0 section 16 defined the various types of adverse events (AEs). All AEs will be recorded, notified, assessed for seriousness, and severity, reported, analysed and managed in accordance with GCP (see Protocol Section 12.2). The reporting and notification of serious adverse events (SAEs) and of SARs that are related and unexpected, are specified in section 12.2 of the protocol.

AEs, SAEs and/or SARs over the study will be summarised by treatment and biomarker.

8. Additional analysis

8.1 Additional analysis

Adjustments will be made in the analysis for baseline variables that are imbalanced between treatments groups, regardless of the significance of the statistical test.

9. Presentation of analysis

TSC meetings will be carried out every year during the recruitment period of the trial. The Trial Statistician and Trial Manager will produce the open reports. A template of the TSC report is saved in Section 22 of the trial TMF.

The study analyses will take place following database lock and will follow the SAP. The final report will be submitted to the Ethics committee.

References

1. Smolen JS, et al., EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2022 update. Ann Rheum Dis. 2023 Jan;82(1):3-18. doi: 10.1136/ard-2022-223356. Epub 2022 Nov 10. Erratum in: Ann Rheum Dis. 2023 Mar;82(3):e76. doi: 10.1136/ard-2022-223356corr1. PMID: 36357155.

2. S. Kelly et al., Ann. Rheum. Dis. 74, 611–617 (2015), doi:10.1136/annrheumdis-2013-204603.

3. Humby, F. et al. Synovial cellular and molecular signatures stratify clinical response to csDMARD therapy and predict radiographic progression in early rheumatoid arthritis patients. Ann. Rheum. Dis. 78, 761–772 (2019).

4. Lewis, M. J. et al. Molecular Portraits of Early Rheumatoid Arthritis Identify Clinical and Treatment Response Phenotypes. Cell Rep. 28, (2019).

5. Lliso-Ribera, G. et al. Synovial tissue signatures enhance clinical classification and prognostic/treatment response algorithms in early inflammatory arthritis and predict requirement for subsequent biological therapy: results from the pathobiology of early arthritis cohort (PEA. Ann. Rheum. Dis. 78, 1642–1652 (2019).

6. Humby, F. et al. Rituximab versus tocilizumab in anti-TNF inadequate responder patients with rheumatoid arthritis (R4RA): 16-week outcomes of a stratified, biopsy-driven, multicentre, open-label, phase 4 randomised controlled trial. Lancet 397, 305–317 (2021).

7. van Riel PL, van Gestel AM. Clinical outcome measures in rheumatoid arthritis. *Ann Rheum Dis*. 2000;59 Suppl 1:i28–i31.

8. Smolen JS, Aletaha D. Scores for all seasons: SDAI and CDAI. *Clin Exp Rheumatol*. 2014;32(5 Suppl 85):S-75–79.

9. Wells G, Becker JC, Teng J, et al. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. *Ann Rheum Dis*. 2009;68(6):954–960.

10. Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. *Health Qual Life Outcomes*. 2003;1:20.

11. Strand V, Singh JA. Improved health-related quality of life with effective disease-modifying antirheumatic drugs: evidence from randomized controlled trials. *Am J Manag Care*. 2008;14(4):239–253.

12. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the functional assessment of chronic illness therapy fatigue scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol*. 2005;32(5):811–819.

13. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991;14(6): 540-545.

14. N. Devlin et al. (eds.), Value Sets for EQ-5D-5L 2022

Appendix A: Deviations from the SAP

This report below follows the statistical analysis plan dated 5th August 2024 apart from following:

Section of report not following SAP	Reason
<insert section >	<insert, e.g. exploratory analyses request by TMG>