Statistical Analysis Plan

GLORIA

The **G**lucocorticoid **L**ow-dose **O**utcome in **R**heumato**I**d **A**rthritis Study
Comparing the cost-effectiveness and safety of additional low-dose glucocorticoid in treatment strategies for elderly patients with rheumatoid arthritis

Sponsor: VU University Medical Center
De Boelelaan 1117
1081 HV Amsterdam
The Netherlands

Author: VU University Medical Center
Department of Epidemiology and Biostatistics &
Amsterdam Rheumatology and immunology **C**enter, location VU University Medical Center
Amsterdam, Netherlands

Date: March 2, 2021

Version: 1.0

Sign-off signatures

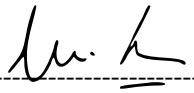
Author: Linda Hartman
PhD student

---------- Signature ---------- Date

Peer review: Leonie Middelink
Project Operational Lead

---------- Signature ---------- Date

Approved by: Maarten Boers
Project leader and Principal Investigator


02-Mar-2021
---------- Signature ---------- Date

Approved by: Johannes Berkhof
Statistician

---------- Signature ---------- Date

Table of contents

List of abbreviations and definitions

aCCP	Anti-CCP
ACR	American College of Rheumatology
AE	Adverse event
AESI	Adverse event of special interest
CI	Confidence Interval
CRP	C-reactive protein
DAS	Disease activity score
DEXA	Dual-energy X-ray absorptiometry
DMARD	Disease-modifying antirheumatic drugs
ESR	Erythrocyte sedimentation rate
EULAR	European League Against Rheumatism
EQ-5D	Euro-QoL in 5 dimensions
GC	Glucocorticoid
GCP	Good Clinical Practice
GLORIA	The Glucocorticoid Low-dose Outcome in RheumatoId Arthritis Study
GRADE	Grades of Recommendations, Assessment, Development, and Evaluation
GTI	Glucocorticoid Toxicity Index
HAQ	Health Assessment Questionnaire
ITT	Intent-to-treat
LRT	Log-likelihood-ratio test
MAR	Missing at random
MCAR	Missing completely at random
MDA	Minimal disease activity
MICE	Multiple imputation by chained equations
NMAR	Not missing at random
NSAID	Non-steroidal anti-inflammatory drug
PP	Per protocol
QALY	Quality-adjusted life years
QoL	Quality of life
RA	Rheumatoid arthritis
RAID	RA Impact of Disease
RF	Rheumatoid factor
SAE	Serious adverse event
SAF	Safety
SAP	Statistical Analysis Plan
SD	Standard Deviation
SDAI	Simple Disease Activity Index
SF-36	Short Form 36-item Health Survey
WHO-ILAR	World Health Organisation/International League of Associations for Rheumatology

1 Introduction

We present the analysis plan executed on the data collected according to the GLORIA protocol. The analysis plan summarizes the study protocol, describes the statistical methods, data listings and summary tables for the analysis of the data from the GLORIA trial. Any differences with the original protocol will be documented.

The details of the economic evaluation and analysis of the adherence substudy will be described in a separate analysis plan.

1.1 Study population

The study population consists of patients of 65 years or older with rheumatoid arthritis (RA) according to the 1987 or 2010 classification criteria of the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) (Aletaha, 2010), requiring antirheumatic therapy because of inadequate disease control, as evidenced by a 28-joint disease activity score (DAS28) of ≥2.60. Patients were selected from 28 hospitals in Germany, Hungary, Italy, The Netherlands, Portugal, Romania and Slovakia.

1.2 Study design

The GLORIA study is a randomized, double-blind, placebo-controlled pragmatic multicenter clinical trial to assess the effectiveness and safety of a daily dose of 5 mg prednisolone or matching placebo in elderly RA patients. Patients will be randomized into two arms: the experimental arm (receiving prednisolone 5 mg/day) or the control arm (receiving placebo). The trial period is 24 months with 3 additional months during which study medication treatment is tapered. In the first 3 months, co-medication is limited: co-medication can be changed only at baseline or after 3 months, creating a brief period in which the effects of prednisolone can be studied with minimum interference.

Our design emulates the routine care setting: eligibility criteria are minimal, assessments and procedures are tailored to represent standard of care, and concurrent antirheumatic treatment is allowed next to the trial medication with minimal limitations. Thus we expect that almost all elderly RA patients (specifically those with comorbidities) are eligible.

The adherence of all patients in the main GLORIA trial is monitored throughout the study period by an adherence monitoring device loaded into the cap of the drug bottle, and by a count of returned capsules at every clinic visit.

Assessment takes place at varying intervals and includes 7 clinic visits and 3 assessments by telephone. The sequence and duration of all study periods is represented in Figure 1.

Figure 1: Schedule of assessments

	Base-line	2 year treatment period								Follow-up / taper
Visit No	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Month (visit window +/- 2 months)	*0*	*3*	*6*	*9*	*12*	*15*	*18*	*21*	*24*[LV]	*27*
Location	*clinic*	*clinic*	*clinic*	*remote*	*clinic*	*remote*	*clinic*	*remote*	*clinic*	*clinic*
Written informed consent	X									
In- / exclusion criteria	X									
Demographics, education & medical history	X									
Baseline prognostic factors: adherence, health literacy, arthritis helplessness index	X*									
Randomization	X*									
Physical examination	X									
Height	X								X	
Vital signs / weight	X	X	X		X		X		X	X
AE evaluation (inc. surgery, comorbidity) (safety)		X*	X*	X*	X*	X*	X*	X*	X*	X*
Concomitant medication	X	X	X	X	X	X	X	X	X	X
Lab (hematology[1], chemistry[2])	X	X	X		X		X		X	
Lab lipids[3]	X	X			X[3]				X	
Lab CRP, ESR	X	X	X		X		X		X	X
Joint counts (44) (effectiveness)	X								X	
Joint counts (28) (effectiveness)		X	X		X		X			X
DEXA either with VFA, or without VFA plus X-lat spine (thor, lumb)[4,5]	X								X	
X-rays hand and forefeet[4] (safety)	X								X	
Patient global assessment, stiffness duration, stiffness severity	X	X	X	X	X	X	X	X	X	X
Physician global assessment	X	X	X		X		X		X	X
Questionnaires:										
- RAID	X*	X*	X*	X*	X*	X*	X*	X*	X*	X*
- full HAQ-DI (24 items)	X*	X*	X*		X*		X*		X*	X*
- MHAQ-DI (10 items)				X		X		X		
- Cost questionnaire	X*	X*	X*		X*		X*		X*	X*
- EQ-5D	X*	X*	X*	X*	X*	X*	X*	X*	X*	X*
- SF36	X*								X*	
- Patient symptom list	X*								X*	X*
GC or placebo, dispense adherence	X*	X*	X*	X*	X*	X*	X*	X*	X*	

** procedures not belonging to standard of care*

lv End of study visit. This visit also needs to be completed by subjects who early terminated the trial.

Apart from CRP and ESR, the following standard of care lab tests will be done on peripheral blood samples.

Lab results of CRP and ESR no more than 4 weeks old may be used for the baseline visit.

Results of the below tests [1, 2,3] no more than 2 months old may be used for the baseline visit:

1) Hemoglobin, Mean Cell Volume, White blood cell count, White blood cell differential count, Platelet count.

2) Glucose (non-fasting), Creatinine, Alanine Aminotransferase;

3) Total cholesterol, HDL cholesterol. Optional for month 12.

4) For the baseline visit: DEXA, DEXA VFA and X-ray images with a maximum of 6 months before or until 3 months after the baseline visit may be used.

5) If possible a whole body DEXA may be performed.

1.3 Primary and secondary study objectives

1.3.1 Primary objectives

The primary objectives of the GLORIA project are to compare the effect of low-dose GC therapy (5 mg/day), relative to placebo, on time-averaged mean value of the disease activity *AND* the number of patients experiencing at least one AE of Special Interest for two years as co-treatment in elderly RA patients (≥65 years).

1.3.2 Secondary objectives

- To assess study medication adherence through a medication packaging solution; and, in a substudy, test the effectiveness of smart device technology to improve adherence (separate statistical analysis plan, not described here).
- To assess the secondary outcome measures which are described in section 1.4.2.
- Deliver an outcome prediction model for individual patient outcome, to tailor treatment strategies for elderly RA patients with comorbidities
- Deliver data to support:
    - Better guidelines on RA treatment in elderly
    - More accurate information for elderly RA patients, their physicians and researchers
    - Improved strategies for trial design and conduct in the elderly

1.4 Primary and secondary endpoints

1.4.1 Primary endpoints

• To measure benefit, the primary endpoints are:
a) signs and symptoms: the time-averaged mean value (estimated from linear mixed models) of the DAS28;
b) damage score: change from baseline after 2-years in total Sharp/van der Heijde damage score of hands and forefeet radiographs.

• To measure safety, the primary endpoint is the total number of patients experiencing at least one adverse event of Special Interest (AESI). An AESI is defined as either a serious adverse event (SAE) according to the GCP definition, or the occurrence of one of the following AEs ('other AESI'):

- Any AE (except loss of efficacy, worsening of disease) that leads to the definite cessation of the trial medication;
- A cardiovascular event (myocardial infarction, cerebrovascular event, peripheral arterial vascular event);
    - Anything falling outside this definition is not an AESI.
      So venous hemorrhoids, venous ulcus cruris, thrombosis, pulmonary embolism, subdural hematoma, and traumatic hemorrhagic brain lesions are not coded as AESI. Also cardiac rhythm disturbances, valve lesions, cardiac insufficiency, cardiomyopathy, and encephalopathy, are not coded as AESI unless part of an acute ischemic episode.
- Newly occurring hypertension requiring drug treatment;

- Newly occurring diabetes mellitus requiring drug treatment;
- Symptomatic bone fracture requiring treatment;
    - AESI is coded as 'no' if the vertebral fracture was asymptomatic, and 'yes' for all nonvertebral fractures, unless expressly indicated that the fracture was asymptomatic.
      Note that vertebral fracture is also captured as secondary outcome through DEXA or spine radiographs.
- Infection requiring antibiotic treatment;
    - Any specific treatment aimed at microorganisms to treat an infection is included: includes topical, antiviral, antifungal, antiparasitic treatment.
    - (Surgical) procedures with concomitant antibiotics: not (initially) an infection, so not AESI; AESI only when the procedure was done for a primary infection (e.g. abscess). Surgical complications treated with antibiotics are also not coded as AESI, because it is usually impossible to determine whether the antibiotic was given as prophylaxis or as treatment.
      Note: almost any surgery is of necessity (hospital admission) already an SAE.
- Newly occurring cataract or glaucoma.
    - Cataract and glaucoma are assumed to be bilateral.
      A medical history of the disease in one eye counts for both, so no AESI during the trial for procedures on the contralateral eye. We count bilateral procedures for new disease as one (not two) AESI.

AEs can be reported spontaneously by patients, on prompting at clinic visits or by telephone interviews, and by the treating physician. The AEs specified above will be recorded with special care in a separate procedure. This procedure includes adjudication through the collection of evidence. This can be a declaration of the treating physician (in case of start of treatment, or a copy of a letter confirming the event, kept on file in the center.

The safety database uses the regular term and definition to describe SAE. The electronic Case Record Form and study database uses AESI to describe events that match one of the other bullets, regardless of whether the event is an SAE. So there an event can be labeled SAE (y/n) and separately AESI (y/n). In the primary analysis patients will be counted who have at least one SAE or AESI as scored in the safety database ('other AESI'), to match with the definition of primary endpoint specified above.
From here, the primary endpoint for safety is defined as 'SAE or other AESI' because one AE can be labeled as a SAE as well as an AESI.

- Cost-effectiveness and cost-utility: more details are described in the separate analysis plan for the cost analyses

- Medication adherence: adherence to trial drug is measured through the e-communicative packaging solution as the count of days in which the bottle is opened on the appropriate days, as measured by the adherence tool. The tool (cap) is collected during the study visits at 3, 6, 12, 18, 24 and 27 months. Each cap is sent back to the provider and the data are read out centrally.

1.4.2 Secondary endpoints

Benefit

- WHO-ILAR core set of RA outcome measures, including patient and physician global assessment of disease activity, joint counts (swollen joints and tender joints), acute phase reactants (C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR)) at 0, 3, 6, 12, 18, 24 and 27 months, and radiographs of hands and forefeet at 0 and 2 years.
  Remaining core set measures: pain and fatigue are part of the RAID questionnaire (see below), and physical function is assessed by the (m)HAQ.
- DAS44 – at 0 and 24 months
- DAS28 – at 3, 6, 12, 18, and 27 months
- Severity and duration of morning stiffness - at every study visit
- SF36 – The Short Form 36-item Health Survey, a questionnaire about quality of life (QoL) - at 0 and 24 months
- 57-symptom list – at 0 and 24 months
- RA Impact of Disease (RAID) tool – The RAID is a validated questionnaire assessing the seven most important domains of impact of RA on the patients: pain, functional disability assessment, fatigue, sleep, physical well-being, emotional well-being, coping. – at every study visit
- (Modified) Health Assessment Questionnaire ((m)HAQ) at every study visit
- RAPID3, calculated with the patient global assessment of disease activity, (m)HAQ and the question about pain from the RAID questionnaire, at every study visit
- Cost questionnaire at 0, 3, 6, 12, 18, 24, 27 months, including
  - Activity limitation
  - Work disability (for those holding a paid job)
- Utility/Quality-adjusted life years (QALY): Euro-QoL in 5 dimensions (EQ-5D) - at every study visit

Harm

- Vital signs (heart rate and blood pressure), height, weight – at 0, 3, 6, 12, 18, 24, 27 months
- Bone mass assessed by Dual-energy X-ray Absorptiometry (DEXA) - at 0 and 24 months
- Vertebral Fracture Analysis (by DEXA OR lateral radiograph of thoracic and lumbar spine) - at 0 and 24 months
- Discontinuation of study drug with reason: this includes patients in whom treatment with prednisolone becomes clinically indicated or those with unacceptable side effects attributable to study medication
- Change of antirheumatic treatment, with reason
- Intensification of treatment for existing comorbidity, e.g. hypertension or diabetes
- Joint replacement surgery
- AEs, SAEs, other AESI at every study visit
- Concomitant medication at every study visit

Data collection is in agreement with industry standard (i.e. Meddra terminology etc.).

GC harm/benefit balance

- Blood samples will be collected during the study as part of the standard of care of RA patients, of which non fasting blood samples at 0, 3, 6, 12, 18, 24 and 27 months. No additional blood samples will be taken for this study. In specific cases, extra lipid analysis will be performed on the blood samples collected as part of the standard of care.

Medication adherence

- Adherence to the trial drug is measured through the e-communicative packaging solution as the count of days in which the bottle is opened on the appropriate days (pill count), as measured by the adherence tool at 3, 6, 12, 18, 24 and 27 months

Other

- Demographics, education and medical history – at baseline
- Baseline prognostic factors: Morisky medication adherence scale (MMAS-8), health literacy, arthritis helplessness index questionnaires – at baseline
- Cost-effectiveness and cost-utility: estimate of costs of treatment and monitoring
- Patients retained on study drug over time: survival analysis

2 Analysis populations

For the full trial period, we will perform analyses of the primary and secondary outcome variables on the intent-to-treat (ITT) population, with ad-hoc sensitivity analyses excluding selected patients based on blinded data review. The ITT principle asserts the effect of a treatment policy (that is, the planned treatment regimen), rather than the actual treatment given (i.e., it is independent of treatment adherence etc.). Accordingly, participants allocated to a treatment group (GC and placebo, respectively) should be followed up, assessed and analysed as members of that group, irrespective of their adherence to the planned course of treatment (i.e., independent of withdrawals and cross-over phenomena). Primary safety analyses will be repeated in the safety (SAF) population.

2.1 Intent-to-treat population

The intent-to-treat population is defined for efficacy as all patients randomized into the study who took the study medication (prednisolone or placebo) for at least one day and who have at least one baseline and one follow-up assessment. Patients will be analyzed in the group they were randomized to.

2.2 Per-protocol population, also excluding initial changes in antirheumatic medication

For the 3-month efficacy assessment: this population consists of all patients who during the first 3 months:

- received the study drug;
- had at least 80% adherence to treatment as determined by tablet count;
- completed the baseline and 3-month visit;
- have a DAS28ESR value available on both visits; or a DAS28CRP value on both visits;
- did not start, stop or change the dose of concurrent antirheumatic drugs, or receive glucocorticoid (oral or injection) at baseline or in the period up to the 3-month visit;
- did not have any serious protocol violations as determined in section 3.

The report from the blinded data review and the list of patients excluded from this population have finalized before the database is locked and are attached to this plan (appendix C).

2.3 Safety population

All patients who took at least one capsule of study medication belong to the safety population.

3 Blinded review of the data

Multiple blinded data review meetings with the principal investigator, medical monitor, study lead, trial statistician, a representative of data management and a member of the scientific advisory committee will be held before breaking the blind. During the blinded data review the protocol violations and their impact on the study outcomes will be discussed. The decisions taken during the blinded data review will be finalized and signed by the participants of this meeting before breaking the blind. The final report of these meetings are attached as appendix C.

4 Statistical and analytical methods

4.1 General statistical considerations

All continuous variables will be summarized with the following descriptive statistics: number of patients (n), mean, standard deviation (SD). If the data are non-normally distributed, median, minimum (min.) and maximum (max.) will be reported. For categorical variables, the absolute (N) and relative frequency (%) will be reported.

4.2 Missing data handling

Handling of missing DAS28 (disease activity) outcomes (applies to DAS28-ESR and DAS28-CRP)
Step 1. Imputation of missing DAS28 if only the component 'patient global assessment of disease activity' is missing
If the missing DAS28 score is from an intermediate (i.e. not baseline or end) visit and only the patient global assessment of disease activity is missing, it will be replaced by the mean of the patient global score immediately before and after this visit, and the DAS28 score will be calculated with this result. If one or both of these are also missing the value will remain missing because we have no information to impute a value.
Step 2. All remaining missing DAS28
In principle, DAS28 outcomes may be incomplete for two reasons:

1. Non-monotone missingness, for example because a patient skips a clinic visit/measurement occasion but is observed again at later points in time
2. Monotone missingness in case of premature discontinuation of study medication

For the primary analysis, we use single imputation by chained equations using the MICE package (van Buuren 2012) to impute missing values according to the following rules. At each study time point we observe or do not observe patient's DAS28, DAS28-CRP, and RAPID-3 score. We impute missing values on these variables per measurement occasion, if at least one measurement is present. If all three measurements are missing, we do not impute any value. These cases usually occur due to discontinuation of the study. The main analysis model (section 4.10) is a mixed effects longitudinal model which gives consistent estimates under the assumption that the remaining missing DAS28 outcomes are missing at random (Little and Rubin 2002).

Sensitivity analysis: Nonresponder imputation
Finally, in case of monotone missingness we investigate the possibility that the missingness occurs not at random (NMAR) with robustness analyses (nonresponder imputation). This imputation method recodes the outcomes of all complete units observed at the end of the study period (j=m), Y_m, as follows:

$$\tilde{Y}_m = \begin{cases} 1 & iff\ Y_m - Y_1 < \text{-}0.6 \\ 0 & iff\ Y_m - Y_1 \geq \text{-}0.6 \end{cases}$$

where 0.6 denotes a minimum effect. Incomplete units receive $\tilde{Y}_m = 0$ representing the conservative hypothesis that all incomplete cases would not have achieved a minimum treatment effect.

Subsequently, we test whether the proportions of responders ($\tilde{Y}_m = 1$) differs significantly between treatment and placebo with a model for difference in proportions akin to model 3 in section 3.3.

Handling of missing damage scores
Step 1. Handling of assessments with partially missing data
Joints with signs of surgery (e.g. prosthesis, arthrodesis) are not scored. Also, when assessments at both time points have been done, but groups of joints are missing from one time point, the scores from that group are also set to missing at the other time point. For example if the end assessment includes only hand films (feet films missing), the initial assessment of the feet will be set to missing. Likewise, if surgery is done on a joint during the trial, the initial assessment of that joint will also be set to missing.

Step 2. All remaining missing damage assessments
We may distinguish three missing data patterns in the damage scores at study onset and study end-point:

1. Damage score is not observed at study onset but the endpoint is observed,
2. Damage score is observed at study onset but the endpoint is unobserved,
3. Damage score is neither observed at study onset nor endpoint.

The three missing data patterns are addressed as follows.

For the primary analysis, we use the algorithm 'MICE' (multiple imputation by chained equations) to impute missing measurements (van Buuren, 2012). This procedure assumes missing at random (MAR) data. The imputation models use data available at baseline (damage score at baseline, disease duration, rheumatoid factor positivity, ACPA positivity, DAS28-ESR) and the endpoint measurement of damage. Study center is not taken into account in this imputation model in view of the large number of cases with missing endpoints.

As a sensitivity analysis, we conduct a complete case analysis deleting all cases with (remaining) missing onset outcomes or missing endpoint outcomes list-wise. This analysis thus excludes cases with patterns 1, 2 and 3. This analysis assumes these outcomes are missing completely at random, MCAR (Little & Rubin, 2002).

Handling of missing harm outcomes
Harm is verified throughout the trial via remote and onsite checks against the source data. So harm assessment is assumed to be complete.

4.3 Baseline definition

Baseline is defined as the last value immediately before the start of the treatment.

4.4 Definition of study visit and visit window

4.4.1 Study visit

The baseline visit (visit 1, in the clinic) is defined as study month 0. The next visits are at 3 months (visit 2, in clinic), 6 months (visit 3, in clinic), 9 months (visit 4, remote), 12 months (visit 5, in clinic), 15 months (visit 6, remote), 18 months (visit 7, in clinic), 21 months (visit 8, remote), 24 months (visit 9, in clinic), 27 months (10, in clinic). The end of the study is defined as the date of the last study visit.

4.4.2 Visit window

The protocol allowed to minimize or maximize the visit window with two months. For the radiology data (x-rays hand and forefeet and DEXA), the time window for the baseline measurement was 6 months before or 3 months after the baseline visit, and for the 24-months measurement 3 months before or after visit 9. If these are exceeded we will accept images up to 1 year before and 6 months after baseline, and up to 1 years after year 2. Data will be analyzed based on the visit number/month that is reported in the database. The calendar visit dates will not be used.

4.5 Patient disposition

The number and percentages of patients allocated to the prednisolone and placebo group will be summarized for all enrolled patients. The number and percentage of patients completing and withdrawing the study, the reasons of withdrawal, and study duration will be reported in a table overall and by treatment.

4.6 Demographics and baseline characteristics

Demographics and baseline characteristics, including birth month and year, sex, smoking , alcohol use, the possession of a smartphone, education level, rheumatoid factor (RF) status, anti-CCP (aCCP) status, duration of RA, evidence of structural joint damage, previous related surgical procedures, previous glucocorticoid use, previous NSAID use, previous DMARD use, previous biological use, number of comorbidities will be reported in a table by treatment group and overall.

4.7 Medical history

Number and percentages of patients with a history of diseases, abnormalities or surgery will be summarized by treatment group and overall. The reported term, preferred term, system organ class, and year of resolution or ongoing will be listed. The MedDRA coding dictionary, version 21.0, will be used to code the medical history.

4.8 Benefit analyses

4.8.1 Primary analyses

4.8.1.1 DAS28

The following primary benefit hypotheses will be tested for the DAS28 score:

1a. Primary benefit hypotheses:

$H_{0,1a}$: (Null) Prednisolone intake over a period of 24 months leads on average to no difference or a higher score in DAS28 than placebo intake.

$H_{1,1a}$: (Alternate) Prednisolone intake over a period of 24 months leads on average to a lower score in DAS28 than placebo intake.

$$H_{0,1a}: \mu_T(t = 24) - \mu_C(t = 24) \geq 0$$

$$H_{1,1a}: \mu_T(t = 24) - \mu_C(t = 24) < 0$$

$\mu_T(t = 24)$ is the mean of DAS28 over 24 months in the treatment group. $\mu_C(t = 24)$ is the mean of DAS28 over 24 months in the control group.

4.8.1.2 Damage score/progression

Damage progression is analyzed as damage at end, with baseline damage as covariate (see model, section 4.10).

The following primary benefit hypothesis will be tested for damage:

2a. Primary benefit hypothesis:

$H_{0,2a}$: (Null) Prednisolone intake over a period of 24 months leads on average to no difference or a higher damage progression compared to placebo intake.

$H_{1,2a}$: (Alternate) Prednisolone intake over a period of 24 months leads on average to a lower damage progression compared to placebo intake.

$$H_{0,2}: \delta_T(t = 24) - \delta_C(t = 24) \geq 0$$

$$H_{1,2}: \delta_T(t = 24) - \delta_C(t = 24) < 0$$

$\delta_T(t = 24)$ is the mean damage score at 24 months in the treatment group. $\delta_C(t = 24)$ is the mean damage score at 24 months in the control group.

4.8.2 Secondary analyses

The following secondary benefit hypothesis will be tested for the DAS28 in the per-protocol population (see 2.2):

1b. Secondary benefit hypotheses:

$\mathrm{H}_{0,1\mathrm{b}}$: (Null) Prednisolone intake over a period of 3 months leads on average to no difference or a larger score in DAS28 than placebo intake.

$\mathrm{H}_{1,1\mathrm{b}}$: (Alternate) Prednisolone intake over a period of 3 months leads on average to a smaller score in DAS28 than placebo intake.

$$H_{0,1b}\colon \mu_{\mathrm{T}}(t=3) - \mu_{\mathrm{C}}(t=3) \geq 0$$

$$H_{1,1b}\colon \mu_{\mathrm{T}}(t=3) - \mu_{\mathrm{C}}(t=3) < 0$$

4.9 Safety analyses

4.9.1 Primary analyses

The following primary harm hypothesis will be tested:

3a. Primary harm null hypothesis:

$H_{0,3a}$: (Null) The probability of patients receiving prednisolone to encounter at least one SAE or other AESI (as defined in the protocol) in a period of 24 months is equal to or less than the probability of patients receiving placebo.

$\mathrm{H}_{1,3\mathrm{a}}$: (Alternate) The probability of patients receiving prednisolone to encounter at least one SAE or other AESI in a period of 24 months is greater than the probability of patients receiving placebo.

$$H_{0,3a}\colon \pi_{\mathrm{T}}(t=24)/\pi_{\mathrm{C}}(t=24) \leq 1$$

$$H_{1,3a}\colon \pi_{\mathrm{T}}(t=24)/\pi_{\mathrm{C}}(t=24) > 1$$

$\pi_{\mathrm{T}}(t=24)$ is the probability to encounter an SAE or other AESI over a 24 months period in the treatment group. $\pi_{\mathrm{C}}(t=24)$ is the probability to encounter an SAE or other AESI over a 24 months period in the control group.

4.10 Modeling and estimating

The testing of the hypothesis sets requires estimating the average effect of treatment on benefit (operationalized as DAS28 and progression of joint damage) and harm (operationalized as encountering at least one SAE or other AESI). An observed effect with a probability of occurrence under the (one-sided) null hypothesis (p-value) of 5% or less will be considered significant causing us to reject the null and to accept the alternative hypothesis. The hypotheses and their tests are one-sided in view of pre-existing knowledge for benefit and the likelihood that true benefit will be underestimated in the context of the pragmatic design and potential differential co-intervention (see section 4.11); and for harm because of limited power. This results in a higher chance of declaring benefit, but also a higher chance of declaring harm.

The average effect of treatment on DAS28 (hypotheses 1) and damage progression (hypothesis 2) will be estimated in a mixed effects regression model. The exact form and empirical fit of the model are important determinants of the power of the statistical test of the treatment effect parameter, as described below. Hypotheses 3 require estimation of the difference in probability to encounter a SAE or other AESI between treatment and placebo over 24 months, which will be implemented by estimating the odds ratio of a SAE or other AESI in treatment versus placebo. All analyses will also take into account that the GLORIA study uses a stratified sampling design, where within each of 28 participating centers randomization is executed stratified by earlier exposure to glucocorticoids (yes/no) and start or switch of other anti-rheumatic drugs at the beginning of the trial.

We first describe how hypothesis testing for the GLORIA study will be executed. In this exposition, we assume that the observed data matrix is complete and that the estimated effects are not confounded with treatment changes. We describe methods to address incomplete data (section 4.2) and to assess confounding with treatment changes (section 4.11).
Our general analysis strategy can be summarized as follows:

1. Test hypotheses 1 to 3 with models described in this section and missing-at-random data correction (section 4.2).
2. Execute robustness checks for monotonous not-missing-at-random data pattern with non-responder imputation (NRI; section 4.2).
3. Assess confounding with treatment changes (section 4.11)
4. Formulate rules to interpret the overall trial results (section 4.12)

4.10.1 Model and statistical test for the effect of treatment on DAS28 (Hypotheses 1)

We first consider the analysis of the DAS28 outcome (benefit). The measurement level of DAS28 may be considered continuous so that we can model outcomes by continuous variable models. In the analysis, we need to solve two problems: first, how to estimate an average effect across the study period of two years, and second, how to take the longitudinal data structure into account. The second aspect poses two sub-problems; first, whether to treat time as continuous or discrete and, second, how to take into account that repeated measurements on the same individual are correlated (standard regression techniques assume uncorrelated observations).

The problem of intra-individual correlated observations will be addressed by mixed-effect regression models. The primary distinction in mixed-effect models for longitudinal data is the treatment of time in the model. In so-called fixed-occasion designs, time takes on a discrete form, whereas in so-called variable-occasion designs time takes on a continuous form (Hox, 2010; Snijders & Bosker, 2011). Fixed occasion designs are characterized by the fact that at given occasions outcomes of all (or most) of subjects are observed. Variable occasion designs are not limited by this restriction and there may be a large number of occasions at which only few subjects need to be observed.

Although the design of the GLORIA study in principle lends itself to both types of model classes we will test hypotheses 1 with a fixed occasion model, because the number of time points at which individuals are observed is limited (up to six) and the model can accommodate non-linear trends in the data well by including occasion specific effects and occasion-treatment interactions (see expectations in figure 2). Continuous time models may also accommodate non-linear trends, but then the specification of the functional form becomes crucial and we do not have strong prior expectations of any particular functional model. We note continuous time models may be considered for secondary analyses.

4.10.1.1 Full model

We model the treatment outcome by a mixed effect model with fixed effects measurement occasion after baseline (d), baseline measurement (Z), treatment condition (T), and random effects across individuals (level 2; random intercept U) and study centers (level 3; random intercept V). Furthermore, we include the strata of the GLORIA trial (S) as fixed effects. Centers could also be modeled as fixed effects, but we have chosen random effects because several centers have included only a very low number of patients. A compound symmetry correlation structure is assumed.

We write down mixed effect model 1 as:

$Y_{ijk} = \beta_{0jk} + \sum_{s=1}^{4} \beta_{1jks} d_{ijks} + \epsilon_{ijk}$ (level-1 model)

$\beta_{0jk} = \delta_{00k} + \delta_{01k} S_{jk1} + \delta_{02k} S_{jk2} + \delta_{03k} T_{jk} + \delta_{04k} Z_{jk} + U_{0jk}$ (level-2 model for intercept)

$\beta_{1jks} = \delta_{10ks}$ (level-2 model for slope)

$\delta_{00k} = \gamma_{000} + V_{00k}$ (level-3 model for intercept)

$\delta_{01k} = \gamma_{010}$ (level-3 model for slope)

$\delta_{02k} = \gamma_{020}$ (level-3 model for slope)

$\delta_{03k} = \gamma_{030}$ (level-3 model for slope)

$\delta_{04k} = \gamma_{040}$ (level-3 model for slope)

$\delta_{10ks} = \gamma_{100s}$ (level-3 model for slope)

with U_{0jk} the individual-level random effect, V_{00k} the center-level random effect, and ϵ_{ijk} the residual.

The indices:

i= 1,…, m : fixed measurement occasion after baseline (in GLORIA, m=5)

j= 1,…,n : individuals

k=1,…,K : study centers (currently, K=28 in GLORIA)

The variables:

Y_{ijk} : DAS28 outcome

d_{ijks} : dummy variables encoding the five measurement occasions after baseline, s=1,…, 4

T_{jk} : treatment indicator, 0 if placebo, 1 if active (prednisolone) treatment

Z_{jk}: Baseline measurement of DAS28

S_{jk1}, S_{jk2}: Stratification factors

The parameters:

γ_{000}: the model intercept

γ_{010}: fixed effect of stratification factor 1

γ_{020}: fixed effect of stratification factor 2

γ_{030}: fixed effect of treatment

γ_{040}: fixed effect of baseline measurement of DAS28

γ_{100s}: effect of time dummy s

And the random effects :

U_{0jk} : random intercept of individual i

V_{00k}: random intercept of study center

ϵ_{ijk}: residual

In hypotheses 1, we test for a difference in mean DAS score across a time of 24 months between treatment and placebo. In model (1), this difference in means is estimated by parameter γ_{030}. For the PP analysis at 3 months, the model becomes two-level (no time) as there is only one assessment post baseline.

4.10.1.2 Testing and exploration of treatment-time interactions

The significance of all parameters in model 1 is evaluated using t-tests with Satterthwaite's method for approximating degrees of freedom (Kuznetsova, Brockhoff, and Christensen, 2017). In particular, hypothesis H1a is tested by evaluating the one-sided p-value of γ_{030} rejecting the one-sided null hypotheses if $\gamma_{030} < 0$ significantly and not rejecting else. The main effect of measurement occasions across all dummy variable d is evaluated jointly by a likelihood ratio test (LRT).

After this assessment, we will explore in a secondary analysis if the interaction between treatment and measurement occasions (time) interaction is significant using a LRT. If this test is significant we will test at which measurement occasions there is a significant difference using t-tests with Bonferroni correction of the p-values.

4.10.2 Model and statistical test for the effect of treatment on damage score (hypothesis 2)

The treatment effect on the damage score at 24 months (hypothesis 2) will be evaluated by the total Sharp/van der Heijde damage score of hands and forefeet. We will use a linear mixed-effects regression model defined similarly to model 1. Contrary to model 1 there are no repeated measurements. Therefore, we model outcomes of individuals nested in centers and include as main effects treatment, design strata, and the baseline measurement of the damage score. A random effect for study center is included. Model 2 is formalized as follows:

$$Y_{jk} = \delta_{0k} + \delta_{1k}S_{jk1} + \delta_{2k}S_{jk2} + \delta_{3k}T_{jk} + \delta_{4k}Z_{jk} + U_{jk} \quad \text{(level-1 model)}$$
$$\delta_{0k} = \gamma_{00} + V_{0k} \quad \text{(level-2 model for intercept)}$$
$$\delta_{1k} = \gamma_{10} \quad \text{(level-2 model for slope)}$$
$$\delta_{2k} = \gamma_{20} \quad \text{(level-2 model for slope)}$$
$$\delta_{3k} = \gamma_{30} \quad \text{(level-2 model for slope)}$$
$$\delta_{4k} = \gamma_{40} \quad \text{(level-2 model for slope)}$$

where now Y_{jk} is the damage score at 24 months and Z_{jk} is the damage score at baseline. Equivalent to section 3.1.1, it can be see that if a one-sided t-test of $\gamma_{30} < 0$ is significant, $H_{0,2}$ is rejected.

We note that the damage progression measurement can be positively skewed, threatening the accuracy (type 1 error rate) of the statistical tests in small samples (and other tests of regression coefficients including t- and F-tests). If we observe a strongly skewed distribution we will apply a bootstrapped test procedure instead of the t-test.

4.10.3 Model and statistical test for the effect of treatment on harm: SAEs or other AESIs (Hypotheses 3)

Contrary to the hypothesis set 1 on the benefit of prednisolone treatment, hypotheses 3 require testing the difference in probabilities of encountering at least one SAE or other AESI over 24 months. On marginal level it would be sufficient to estimate these probabilities after t=24 months and evaluate the significance of their difference. However, to additionally account for the stratification of the GLORIA design, we apply a logistic mixed-effects regression (model 3):

$$P_{jk} = g^{-1}(\delta_{0k} + \delta_{1k}S_{jk1} + \delta_{2k}S_{jk2} + \delta_{3k}T_{jk} + U_{jk})$$

$\delta_{0k} = \gamma_{00} + V_{0k}$ (level-2 model for intercept)
$\delta_{1k} = \gamma_{10}$ (level-2 model for slope)
$\delta_{2k} = \gamma_{20}$ (level-2 model for slope)
$\delta_{3k} = \gamma_{30}$ (level-2 model for slope)
$\delta_{4k} = \gamma_{40}$ (level-2 model for slope)

where now P_{jk} the probability of patient j in center k to encounter at least one SAE or other AESI over 24 months. The function g^{-1} denotes the inverse of the log link. Significance of δ_{3k} will be evaluated with a t-test. Note that model 3 uses the log link as opposed to, for example, the logit link to estimate and test relative risk.

4.11 Evaluation of possible confounded results due to differential co-intervention

The data collected according to the GLORIA protocol are intended to allow a comparison of the effect of prednisolone treatment over placebo in terms of benefit and harm. Given that GLORIA is a pragmatic trial where changes in concomitant treatment (co-interventions for RA, i.e. antirheumatic co-medication) are allowed, an estimation of “pure effects” (i.e. without changes in concomitant treatment) is not possible. When changes in antirheumatic treatment differ between the groups this can result in a confounded estimate of the “pure” effects. Most changes in co-medication can be expected to lower the chance of finding a difference in efficacy between the treatment groups, because negative consequences of being in the placebo group can be countered by changing antirheumatic therapy. For this reason, we also analyze the short-term efficacy of prednisolone in the first three months in a per-protocol population where the chance of confounding is minimized. Note: adverse events leading to stop of study medication are part of the primary harm outcome, and are not covered here.

We will test for potential confounding of the efficacy results. We will not test for confounding of the harm results, because straightforward interpretation is difficult: harm and changes in medication are probably related to each other in a complicated way.

4.11.1 Testing and correcting for differential treatment changes

We plan to check for substantial confounding by comparing the incidence of lasting and substantial changes in co-medication categorized by the reason for change. For practical reasons, we will only consider the first instance of such a change in co-medication. Lasting means we do not consider brief starts or interruptions (i.e., lasting no more than 6 weeks); we also disregard changes in patients stopping the study medication within 3 months following the change.
We will only consider changes that occur in the period starting at 3 months (V2) and 15 months (V6), because the potential for confounding is greatest in this period. Changes before V2 are too early to create confounding, and changes beyond V6 have relatively little effect on the outcome (effect on measurement at V7 and V9). Co-medication includes all disease-modifying antirheumatic drugs (conventional, biologic and targeted), and GC treatment classified as protocol violation and adjudicated in case of oral use.
Oral GC treatment is adjudicated to select occurrences with potential for confounding similar to that of other antirheumatic drugs. Periods of oral treatment up to 3 weeks for RA are equated to one im injection for RA. More than two of either in the period assessed will count as intensification.
Oral GC treatment for comorbidity creates less potential for confounding. More than 4 periods of oral treatment for comorbidity lasting no more than 3 weeks will count as intensification.
Any period of oral therapy longer than 3 weeks for any indication (RA or otherwise) will also count as intensification.

We distinguish 3 reasons for co-medication changes:

1. Lack of efficacy: this includes instances of
    a. co-medication intensification: increase dose of current antirheumatic drug(s), or adding a new drug (co-medication) to the current treatment, including adjudicated GC treatment classified as protocol violation;
    b. co-medication switch (replacing one antirheumatic drug with another);
2. Adverse event: co-medication switch, or stop of an antirheumatic drug and start of another later; sole dose reduction for adverse event is not considered a substantial change;
3. Good efficacy: co-medication taper: decrease dose or stop of current medication(s) without addition of or replacement by another medication.

The remaining patients will be categorized as having no change in co-medication.

For every patient we record and categorize the reason (lack of efficacy, adverse event, good efficacy) for the first change in co-medication in the specified time period. If an adverse event was reported at the moment of change(s) in medication, we will assume that the medication is changed due to adverse events.
Otherwise, if a patient has no reported adverse event(s) at the moment of:

- intensifying or switching the medication: we will assume that the treatment is changed due to lack of efficacy;
- lowering or stopping co-medication: we will assume that the change is due to good efficacy.

Situations that are unclear will be adjudicated by the medical monitor before the blind is broken.

The effect of the co-medication change on the expected DAS28 depends on the reason: in case of lack of efficacy or adverse event, the change is expected to lead to a better outcome: a new drug or

an increased dose of the same drug is expected to work better. In case of good efficacy the change is expected to lead to the same or worse outcome: tapering a drug will most likely increase the DAS28 somewhat and may cause a flare.
Therefore, for both treatment groups we will calculate:

1. the sum of patients experiencing a co-medication change for lack of efficacy, and for adverse events; and
2. the number of patients experiencing a co-medication change for good efficacy.

We will express 1) and 2) as proportion of the total number of patients in each treatment group and compare these between the groups with two Z-tests (each one-sided at $p<0.025$).
The null hypotheses are:
$H_{0,4a}$ *:* The probability of experiencing a co-medication change for lack of efficacy or for adverse events in the placebo group is lower than or equal to the probability in the prednisolone group.
$H_{1,4a}$ *:* The probability of experiencing a co-medication change for lack of efficacy or for adverse events in the placebo group is higher than the probability in the prednisolone group.

$H_{0,4b}$ *:* The probability of experiencing a co-medication change for good efficacy in the placebo group is higher than or equal to the probability in the prednisolone group.

$H_{1,4b}$ *:* The probability of experiencing a co-medication change for good efficacy in the placebo group is lower than the probability in the prednisolone group.

We will assume confounding to the detriment of active treatment if:

1. $H_{0,4a}$ is rejected significantly at $p < 0.025$ OR
2. $H_{0,4a}$ is rejected significantly at $p < 0.025$

Assessing potential for confounding by co-medication changes in the analyses regarding harm of prednisolone is more difficult. Harm and changes in medication are probably related to each other in a complicated way. Therefore, we will not examine this.

4.12 Interpretation rules

The outcomes of benefit (disease activity and damage progression) and harm (number of patients with at least one SAE or other AESI) will be interpreted simultaneously.
The outcomes of benefit will be interpreted as follows (prednisolone group compared to placebo; reference to the specific rejected null hypotheses in parentheses):

1. success:
    a. lower disease activity ($H_{0,1a}$)
       AND lower damage score ($H_{0,2a}$)
    b. lower disease activity OR lower damage score; AND confounding ($H_{0,4a}$ or $H_{0,4b}$)
2. partial success/tradeoff:
    a. lower disease activity OR lower damage score AND NO confounding
    b. NO lower disease AND NO lower damage progression AND confounding
3. failure: NO lower disease activity AND NO lower damage score AND NO confounding.

The outcomes of harm will be interpreted as follows:
1. success: NO significant increase in AEs
2. failure: significant increase in AEs ($H_{0,3a}$)

Combined assessment of benefit and harm

a. Success: success in benefit and harm
b. Failure: failure in benefit and harm
c. Partial success/tradeoff: all other scenarios

For further interpretation of increases in the occurrence of SAE or other AESI in the active treatment group that did not reach significance, we will use the suggestions of the GRADE (Grades of Recommendations, Assessment, Development, and Evaluation) Working Group (Guyatt 2012), upper limit of one-sided 95% confidence interval (CI):

1. > 1.3: GC associated with a trend towards greater occurrence;
2. ≤ 1.2: GC appears to have little effect on the occurrence;
3. > 1.2 and ≤ 1.3: results failed to demonstrate or exclude a greater occurrence.

In the case of a numerical DECREASE in the occurrence of SAE or other AESI in the active treatment group, a more stringent test for significance will be performed ($p<0.025$).
For the interpretation of decreases that did not reach significance:
lower limit of one-sided 95% CI:

1. < 0.7: GC associated with a trend towards reduced occurrence;
2. ≥ 0.8: appears to have little effect on the occurrence;
3. < 0.8 and ≥ 0.7: results failed to demonstrate or exclude a reduced occurrence.

4.13 Analysis of secondary outcome variables

We will estimate and test for the detrimental effects of prednisolone on bone mass and incident fracture rate.

1. Continuous outcomes: bone mass lumbar spine, bone mass hip:
   analyzed according to model 1
2. Discrete outcome: presence of at least one incident compression fracture
   (by VFA or lateral spine films): analyzed according to model 2

Tests will be one-sided at $p<0.05$, with Benjamini-Hochberg adjustment for multiple testing.

We will report the results of the other secondary outcome variables descriptively with means/proportions and 95%CI per treatment group per time point.

3. Variables measured at every clinic visit
   a. Continuous outcomes: patient and physician global assessment of disease activity, joint counts (swollen joints and tender joints), acute phase reactants (C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR)); vital signs, lab (blood) variables; adherence by pill count and by electronic adherence cap.
   b. Discrete outcomes: patients with at least one instance of remission; of minimal disease activity; patients with at least one long-term period (2 consecutive instances 6 months apart) of remission; of minimal disease activity
4. Variables measured at every visit (clinic and remote):
   a. Continuous outcomes: severity and duration of morning stiffness, RA Impact of disease (total score, and components pain and fatigue analyzed separately), RAPID3 index, EuroQoL, m(HAQ)
   b. Discrete outcomes: discontinuations of study drug, of with reason; all changes of antirheumatic treatment, with reason; joint replacement surgery
   c. Expressed as survival curve: discontinuation of study drug over time
5. Variables measured at baseline and end of study:
   a. Continuous outcomes: DAS44, SF36 (domains and summary mental and physical component scores), 57-symptom list;
   b. Discrete outcome: number of incident compression fractures (by VFA or lateral spine films)

6. Variables measured only at baseline:
   a. Continuous/discrete:
      Demographics, education and medical history; Morisky medication adherence scale (MMAS-8), health literacy, arthritis helplessness index questionnaires; number of prevalent vertebral fractures (by VFA or lateral spine films), distribution of % height loss in vertebrae with fractures.
7. Response indices calculated at 3 months in the PP analysis (discrete): ACR20, 50, 70; EULAR response.

5 Interim analyses

No interim analyses were planned. However, we performed a new sample size calculation based on blinded AE incidence rates halfway through the trial (see below).

6 Sample size and power calculations

In the main GLORIA trial, 225 patients per treatment group will be entered (total 450 patients).

In the chosen analysis strategy (see section 4), to detect differences in benefit (disease activity, radiographs) extensive RA trial experience (both for GC and other agents) has shown a sample size of 200 patients per group is amply sufficient. For example, in the CAPRA-2 study that compared modified release prednisone 5mg/d against placebo, the prednisone group had 231 patients, the placebo group 119 patients. The change in DAS28 after 3 months of treatment was –1.15 in the prednisone group, –0.63 in the placebo group; difference –0.52 (SE 0.13, $p<0.001$) (Buttgereit, 2013). However, the true incidence of adverse events (AEs) for GCs is currently unknown. Most relevant data to assess sample size adequacy for this study come from the reported CAMERA-2 trial (Bakker, 2012). This trial randomized 236 early RA patients to tight-control high-dose methotrexate plus 10 mg prednisolone or placebo for two years. Interestingly, 22% of placebo patients compared to only 14% of prednisolone patients reported at least one serious AE or clinical event as defined in our protocol.

The original protocol used a base case expectation for a total of 20% of patients experiencing at least one event over two years in the placebo group and calculated a need for 400 patients in each treatment group, to have about 80% power to detect an increase of 7% (from 20% to 27% events; 90% power for an increase of 9%). However, based on our current experience the sample can be decreased to about 450 instead of 800.

Elaboration:
In December 2018 we had >400 patients in the trial with a mean follow up of 9 months, and we had 84 cases with at least one serious adverse event or an event of special interest (the primary harm outcome). A simple extrapolation leads to an estimated rate of about 49% over 2 years (pooled over the whole blinded trial population, i.e. both treatment groups taken together). We have also performed life table analysis, which accounts for patients stopping prematurely; this analysis leads to an estimated rate of 40% (95% confidence interval: 30%-50%).

Given the above we can now assume with confidence that the base rate of patients with events in the placebo group will be substantially higher than originally predicted. This adds power to the trial: we need only between 400 and 450 patients to detect the originally targeted contrast of 27.5%/20%= relative risk of 1.38 (Table 1). At the expected pooled event rate of 40%, with 400 patients we have 80% power to detect a difference of 12% between placebo and prednisolone: i.e. placebo 34% and prednisolone 46%, a relative risk of 1.35. The results are better with 450 patients, and when the event rate is higher (Table 1). At the moment of submission of the amendment trial recruitment has been closed on December 31, 2018 at a total of 452 patients.

Table 1. Sample size scenarios (one-sided alpha 5%).

expected event rate over 2 years (% unique patients with at least one event)		with power 80%, detectable rate in predn group		Relative Risk	
pooled	placebo	n=400	n=450	n=400	n=450
30%	26%	38%	37%	1,47	1,43
35%	30%	42%	41%	1,40	1,37
40%	**34%**	**46%**	**45%**	**1,35**	**1,32**
45%	39%	52%	52%	1,31	1,31
50%	43%	55%	55%	1,28	1,28

7 Tables, listings and figures

Below are the titles for the planned study tables, figures, and listings.

Table 1 Patient distribution
Table 2.A Patient distribution by center (ITT)
Table 2.B Patient distribution by center (PP)
Table 3.A Demographics and baseline characteristics (ITT)
Table 3.B Demographics and baseline characteristics (PP)
Table 4 Health literacy
Table 5 Morisky medication adherence scale (MMAS-8)
Table 6 Arthritis helplessness index Table 7 .A Medical history (ITT)
Table 7.B Medical history (PP)
Table 8.A DAS28 scores (ITT)
Table 8.B DAS28 scores (PP)
Table 9 Mixed effects model of time-averaged DAS28 (ITT)
Table 10 DAS44 scores (ITT)Table 11.A Swollen joints 28 joint count (ITT)
Table 11.B Swollen joints 28 joint count (PP)
Table 12.A Tender joints 28 joint count (ITT)
Table 12.B Tender joints 28 joint count (PP)
Table 13.A Patient global assessment of disease activity (ITT)
Table 13.B Patient global assessment of disease activity (PP)
Table 14.A Physician global assessment of disease activity (ITT)
Table 14.B Physician global assessment of disease activity (PP)
Table 15.A CRP (ITT)
Table 15.B CRP (PP)
Table 16.A ESR (ITT)
Table 16.B ESR (PP)Table 17 EULAR response (PP)
Table 18 ACR response (PP)
Table 19.A Minimal disease activity (ITT)
Table 19.B Minimal disease activity (PP)
Table 20 Total Sharp/van der Heijde damage score of hands and forefeet radiographs (ITT)
Table 21 Morning stiffness (ITT)
Table 22 SF-36 questionnaire (ITT)
Table 23 RAID questionnaire (ITT)
Table 24 (Modified) Health Assessment Questionnaire ((m)HAQ) (ITT)
Table 25 RAPID3 (ITT)
Table 26 EuroQoL questionnaire (ITT)
Table 27 57-symptom list (ITT)
Table 28.A Laboratory results in SI and original units (ITT)
Table 28.B Laboratory results in SI and original units (SAF)
Table 29.A Vital signs (ITT)
Table 29.B Vital signs (SAF)
Table 30.A Adherence (ITT)
Table 30.B Adherence (PP)

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
Patient distribution	1	Intent-to-treat population	Distribution	Baseline	-	Dichotomous	% (n)
Patient distribution by center	2A	Intent-to-treat population	Distribution	Baseline	Treatment group	Dichotomous	% (n)
Patient distribution by center	2B	Per-protocol population	Distribution	Baseline	Treatment group	Dichotomous	% (n)
Demographics and baseline characteristics	3A + 3B	3A. Intent-to-treat population 3B. Per-protocol population	Age	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Sex	Baseline	Treatment group	Dichotomous	% (n) female
			Smoking status	Baseline	Treatment group	Categorical	Current: % (n) Previous: % (n) Never: % (n)

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Alcohol intake (units per week)	Baseline	Treatment group	Categorical	Every day: % (n) 5-6 days per week: % (n) 3-4 days per week: % (n) 1-2 days per week: % (n) 1-3 days per month: % (n) <1 day per month: % (n) Never: % (n)
			Possession of smartphone	Baseline	Treatment group	Dichotomous	% (n)
			Education level	Baseline	Treatment group	Categorical	% (n) higher education
			Duration of rheumatoid arthritis (months)	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Rheumatoid factor status	Baseline	Treatment group	Dichotomous	% (n), positive
			Anti-CCP status	Baseline	Treatment group	Dichotomous	% (n), positive

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Evidence of structural joint damage	Baseline	Treatment group	Categorical	% (n), yes
			Previous related surgical procedures	Baseline	Treatment group	Categorical	% (n), yes
			Previous glucocorticoid treatment	Baseline	Treatment group	Dichotomous	% (n), yes
			Previous NSAID treatment	Baseline	Treatment group	Dichotomous	% (n), yes
			Previous DMARD treatment	Baseline	Treatment group	Dichotomous	% (n), yes
			Previous biological treatment	Baseline	Treatment group	Dichotomous	% (n), yes
			Number of comorbidities	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			DAS28	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Tender joint count	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Swollen joint count	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			ESR	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			CRP	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Morning stiffness, duration in minutes	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed)
			Morning stiffness severity	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Patient global assessment of disease activity (scale 0-10)	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Physician global assessment of disease activity (scale 0-10)	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Pain (scale 0-10)	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			HAQ (scale 0-3)	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			RAID	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			EQ-5D	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			SF36	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
Health literacy	4	Intent-to-treat population	Health literacy	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
Morisky medication adherence scale (MMAS-8)	5	Intent-to-treat population	Morisky medication adherence scale	Baseline	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
Arthritis helplessness index	6	Intent-to-treat population	Arthritis helplessness index	Baseline	Treatment group	Categorical	-Low: % (n) -Normal: % (n) -High: % (n)
Medical history	7A	Intent-to-treat population	Number of comorbidities per organ system class	Baseline	Treatment group	Continuous	% (n)
			Top 3 of most common comorbidity categories	Baseline			1. % (n) 2. % (n) 3. % (n)
	7B	Per-protocol population	Number of comorbidities per organ system class	Baseline	Treatment group	Continuous	% (n)

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Top 3 of most common comorbidity categories	Baseline	Treatment group	Continuous	1. % (n) 2. % (n) 3. % (n)
DAS28 scores	8A	Intent-to-treat population	DAS28 score	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Change in DAS28 score	0-3 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Change in DAS28 score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
	8B	Per-protocol population	DAS28 score	0, 3 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Change in DAS28 score	0-3 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
Mixed effects model of time-averaged DAS28	9	Intent-to-treat population	DAS28 score	0, 3, 6, 9, 12, 18, 24 months	Treatment group	Continuous	Value, Standard Error, Df, p-value
DAS44 scores	10	Intent-to-treat population	DAS44 score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in DAS44 score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
Swollen joints 28 joint count	11A	Intent-to-treat population	Number of swollen joints	0, 3, 6, 9, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
	11B	Per-protocol population	Number of swollen joints	0, 3 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
Tender joints 28 joint count	12A	Intent-to-treat population	Number of tender joints	0, 3, 6, 9, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
	12B	Per-protocol population	Number of tender joints	0, 3 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
Patient global assessment of disease activity	13A	Intent-to-treat population	Patient global assessment of disease activity	0, 3, 6, 9, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
	13B	Per-protocol population	Patient global assessment of disease activity	0, 3 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
Physician global assessment of disease activity	14A	Intent-to-treat population	Physician global assessment of disease activity	0, 3, 6, 9, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
	14B	Per-protocol population	Physician global assessment of disease activity	0, 3 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
CRP	15A	Intent-to-treat population	CRP	0, 3, 6, 9, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI
	15B	Per-protocol population	CRP	0, 3 months	Treatment group	Continuous	(median, min, max if data are non-normal distributed), n
ESR	16A	Intent-to-treat population	ESR	0, 3, 6, 9, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI
	16B	Per-protocol population	ESR	0, 3 months	Treatment group	Continuous	(median, min, max if data are non-normal distributed), n
EULAR response	17	Per-protocol population	EULAR response	3 months	Treatment group	Categorical	Good: % (n) Moderate: % (n) None: % (n)
American College of Rheumatology (ACR) response	18	Per-protocol population	ACR20 response	3 months	Treatment group	Dichotomous	% (n), yes
			ACR50 response	3 months	Treatment group	Dichotomous	% (n), yes
			ACR70 response	3 months	Treatment group	Dichotomous	% (n), yes
Minimal disease activity	19A	Intent-to-treat population	Minimal disease activity	3, 6, 12, 18, 24months	Treatment group	Dichotomous	% (n), yes

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Boolean remission episode	3, 6, 12, 18, 24 months	Treatment group	Dichotomous	% (n), yes
			Long-term Boolean remission	12, 18, 24 months	Treatment group	Dichotomous	% (n), yes
			Simple Disease Activity Index (SDAI)	3, 6, 12, 18, 24months	Treatment group	Dichotomous	% (n), yes
	19B	Per-protocol population	Minimal disease activity	3 months	Treatment group	Dichotomous	% (n), yes
			Boolean remission episode	3 months	Treatment group	Dichotomous	% (n), yes
			Simple Disease Activity Index (SDAI)	3 months	Treatment group	Dichotomous	% (n), yes
Total Sharp/van der Heijde damage score of hands and forefeet radiographs	20	Intent-to-treat population	Sharp/van der Heijde damage score	0, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
			Change in Sharp/van der Heijde damage score	0-24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
Morning stiffness	21	Intent-to-treat population	Duration (minutes) of morning stiffness	0, 3, 6, 9, 12, 15, 18, 21, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
		Intent-to-treat population	Severity of morning stiffness	0, 3, 6, 9, 12, 15, 18, 21, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
SF-36	22	Intent-to-treat population	SF-36 physical functioning score	0 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 physical functioning score	0-24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	SF-36 role-physical score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 role-physical score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
		Intent-to-treat population	SF-36 bodily pain score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 bodily pain score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	SF-36 general health score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 general health score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	SF-36 vitality score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 vitality score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
		Intent-to-treat population	SF-36 social functioning score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 social functioning score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	SF-36 role-emotional score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 role-emotional score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	SF-36 mental health score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 mental health score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
		Intent-to-treat population	SF-36 physical component summary score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 physical component summary score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	SF-36 mental component summary score	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in SF-36 mental component summary score	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
RAID	23	Intent-to-treat population	RAID total score	0, 3, 6, 9, 12, 15, 18, 21, 24months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			RAID pain score	0, 3, 6, 9, 12, 15, 18, 21, 24months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
			RAID fatigue score	0, 3, 6, 9, 12, 15, 18, 21, 24months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
(m)HAQ	24	Intent-to-treat population	(m)HAQ score	0, 3, 6, 9, 12, 15, 18, 21, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
RAPID3	25	Intent-to-treat population	RAPID3 index score	0, 3, 6, 9, 12, 15, 18, 21, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
EuroQoL	26	Intent-to-treat population	EQ-5D score	0, 3, 6, 9, 12, 15, 18, 21, 24months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
57-symptom list	27	Intent-to-treat population	Number of symptoms	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
			Change in number of symptoms	0-24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
Laboratory results	28A+28B	28A. Intent-to-treat population 28B. Safety population	Hemoglobin SI unit (unit: mmol/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Hemoglobin original unit (unit: mmol/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Hemoglobin original unit (unit: g/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Hemoglobin original unit (unit: g/dl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Mean cell volume SI unit (unit: fl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			White blood cells SI unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			White blood cells original unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			White blood cells original unit (unit: /nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			White blood cells original unit (unit: /ul)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			White blood cells original unit (unit: c/nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			White blood cells original unit (unit: G/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			White blood cells original unit (unit: Gpt/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			White blood cells original unit (unit: n/nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			White blood cells original unit (unit: x10^3mm/mm^3)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Neutrophils SI unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Neutrophils relatively (unit: %)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Neutrophils original unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Neutrophils original unit (unit: x10^3mm/mm^3)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Neutrophils original unit (unit: /nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Neutrophils original unit (unit: G/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Neutrophils original unit (unit: c/nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Neutrophils original unit (unit: /uL)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Basophils SI unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Basophils relatively (unit: %)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Basophils original unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Basophils original unit (unit: /nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Basophils original unit (unit: /uL)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Basophils original unit (unit: G/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Basophils original unit (unit: x10^3mm/mm^3)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Eosinophils SI unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Eosinophils relatively (unit: %)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Eosinophils original unit (unit: /nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Eosinophils original unit (unit: /uL)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Eosinophils original unit (unit: G/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Eosinophils original unit (unit: x10^3mm/mm^3)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Lymphocytes SI unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Lymphocytes relatively (unit: %)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Lymphocytes original unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Lymphocytes original unit (unit: /nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Lymphocytes original unit (unit: /uL)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Lymphocytes original unit (unit: c/nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Lymphocytes original unit (unit: G/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Lymphocytes (unit: x10^3mm/mm^3)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Monocytes SI unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Monocytes relatively (unit: %)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Monocytes original unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Monocytes original unit (unit: /nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Monocytes original unit (unit: /uL)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Monocytes original unit (unit: G/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Monocytes original unit (unit: x10^3mm/mm^3)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Platelets SI unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Platelets original unit (unit: x10^9/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Platelets original unit (unit: /nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Platelets original unit (unit: /uL)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Platelets original unit (unit: c/nl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Platelets original unit (unit: G/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Platelets original unit (unit: Gpt/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Platelets original unit (unit: 10^3mm/mm^3)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Glucose SI unit (unit: mmol/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Glucose original unit (unit: mmol/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Glucose original unit (unit: mg/dl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Creatinine SI unit (unit: mcmol/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Creatinine original unit (unit: mcmol/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Creatinine original unit (unit: mg/dl)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Creatinine original unit (unit: mmol/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			ALAT SI unit (unit: U/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			ALAT original unit (unit: U/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			ALAT original unit (unit: ukat/L)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			ALAT original unit (unit: umol/s*l)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Cholesterol SI unit (unit: mmol/L)	0, 3, 12, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Cholesterol original unit (unit: mmol/L)	0, 3, 12, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Cholesterol original unit (unit: mg/dl)	0, 3, 12, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
		-	HDL cholesterol SI unit (unit: mmol/L)	0, 3, 12, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			HDL cholesterol original unit (unit: mmol/L)	0, 3, 12, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			HDL cholesterol original unit (unit: mg/dl)	0, 3, 12, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			CRP SI unit (unit: mg/L)	0, 3, 6, 12, 18, 24, 27 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			ESR SI unit (unit: mm/h)	0, 3, 6, 12, 18, 24, 27 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
Vital signs	29A + 29B	29A. Intent-to-treat population 29B. Safety population	Weight (kg)	0, 3, 6, 12, 18, 24months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
			Height (cm)	0, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Systolic blood pressure (mmHg)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Diastolic blood pressure (mmHg)	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n
			Heart rate per minute	0, 3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean change from baseline (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
Adherence	30A	Intent-to-treat population	% adherence according to pill count	3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	% adherence according to electronic caps	3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
	30B	Per-protocol population	% adherence according to pill count	3 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Per-protocol population	% adherence according to electronic caps	3 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
	30C	Safety population	% adherence according to pill count	3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
		Safety population	% adherence according to electronic caps	3, 6, 12, 18, 24 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
Concomitant medications summary	31	Intent-to-treat population	Concomitant medication summary	Baseline	Treatment group	Dichotomous/continuous	-DMARDs: % (n) -NSAIDs: % (n) -Biologicals: % (n) -Other anti-rheumatic medication: % (n) -Other medication: Mean number (SD) (median, min, max if data are non-normal distributed) -Top 3 of most common concomitant medication class: 1. % (n) 2. % (n) 3. % (n)

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
Bone mass	32	Intent-to-treat population	Bone mass lumber spine	0, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in bone mass lumbar spine	0-24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Bone mass hip	0, 24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Change in bone mass hip	0-24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
		Intent-to-treat population	Number of patients with at least one incident compression fracture (by VFA or lateral spine films)	24 months	Treatment group	Dichotomous	% (n), yes

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
		Intent-to-treat population	Number of incident compression fractures (by VFA or lateral spine films)	24 months	Treatment group	Continuous	Mean (SD) + 95% CI (median, min, max if data are non-normal distributed), n
Number of prevalent vertebral fractures (by VFA or lateral spine films)	33	Intent-to-treat population	Number of prevalent vertebral fractures	0 months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
Distribution of % height loss in vertebrae with fractures	34	Intent-to-treat population	Distribution of % height loss in vertebrae with fractures	0- months	Treatment group	Continuous	Mean (SD) (median, min, max if data are non-normal distributed), n
Joint replacement surgery	35	Intent-to-treat population	Number of joint replacement surgeries	0, 3, 6, 9, 12, 15, 18, 21, 24 months	Treatment group	Continuous	n (%)

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
AE, SAE, other AE of special interest summary	36A	Intent-to-treat population	Number of AEs, SAEs, other AEs of special interest	3, 24 months	Treatment group	Continuous	-Mean (SD) (median, min, max if data are non-normal distributed) -Top 3 of most common AE categories: 1. % (n) 2. % (n) 3. % (n)
	36B	Safety population	Number of AEs, SAEs, other AEs of special interest	3, 24 months	Treatment group	Continuous	-Mean (SD) (median, min, max if data are non-normal distributed) -Top 3 of most common AE categories: 1. % (n) 2. % (n) 3. % (n)
AEs, SAEs, other AEs of special interest	36C	Intent-to-treat population	Number of AE, SAE, other AE of special interest per organ system	3, 24 months	Treatment group	Categorical	-Category 1: n (%) -Category 2: (n%) -Category 3: n (%) Etc.

Table title	Number	Population	Endpoint	Time points or how to conglomerate	Covariates or subgroups	Type of variable	Summary statistics
	36D	Safety population	Number of AE, SAE, other AE of special interest per organ system	3, 24 months	Treatment group	Categorical	-Category 1: n (%) -Category 2: (n%) -Category 3: n (%) Etc.
Presence of SAE or other AE of special interest	37	Intent-to-treat population	Presence of SAE or other AE of special interest	24 months	Treatment group	Dichotomous	-Yes: n(%) -No: n(%)

Listing 1	Patient disposition
Listing 2	Patients excluded from the analysis populations
Listing 3	Protocol violations
Listing 4	Visit windows
Listing 5	Study completion including reason early withdrawal
Listing 6	Demographics and baseline characteristics
Listing 7	History of glaucoma and cataract
Listing 8	RA history
Listing 9.A	Medical history
Listing 9.B	Medical history with MedDRA Coding
Listing 10.A	Treatment not given according to protocol - interrupted
Listing 10.B	Treatment not given according to protocol – discontinued
Listing 11	Change of antirheumatic treatment, with reason
Listing 12	Intensification of treatment for existing comorbidity, e.g. hypertension or diabetes
Listing 13	DAS28 scores, DAS44 scores and EULAR response
Listing 14	Swollen and tender joints
Listing 15.A	Patient global assessment of disease
Listing 15.B	Physician global assessment of disease
Listing 16	ACR 20 response
Listing 17	Minimal disease activity
Listing 18	Total Sharp/van der Heijde damage score of hands and forefeet radiographs
Listing 19	Morning stiffness
Listing 20.A	SF-36 questionnaire
Listing 20.B	SF-36 scores
Listing 21.A	(m)HAQ questionnaire
Listing 21.B	(m)HAQ scores
Listing 22.A	RAID questionnaire
Listing 22.B	RAID scores
Listing 23.A	EQ-5D questionnaire
Listing 23.B	EQ-5D scores
Listing 24	Patient symptom list questionnaire
Listing 25.A	Laboratory – Hematology and chemistry
Listing 25.B	Laboratory – Lipids
Listing 25.C	Laboratory – CRP and ESR
Listing 26	Vital signs
Listing 27	Physical examination
Listing 28	Previous medication
Listing 29	Concomitant medications
Listing 30.A	Adverse events
Listing 30.B	Serious adverse events
Listing 30.C	Other adverse events of special interest
Listing 31	Death

8 Statistical software

The statistical programming and analysis platform R will be used for the analysis (R Core Team, 2020).

SPSS version 26 or higher will be used for the descriptive statistics. Graphpad Prism version 9.0.2 or higher will be used for the figures.

9 References

Aletaha D, Neogi T, Silman AJ, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010 Sep;62(9):2569-81.

Bakker MF, Jacobs JW, Welsing PM, et al. Low-dose prednisone inclusion in a methotrexate-based, tight control strategy for early rheumatoid arthritis: a randomized trial. Annals of Internal Medicine 2012;156:329-39.

Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 1995;57:289-300.

Buttgereit F, Mehta D, Kirwan J, Szechinski J, Boers M, Alten RE, et al. Low-dose prednisone chronotherapy for rheumatoid arthritis: a randomised clinical trial (CAPRA-2). Ann Rheum Dis 2013;72:204-10.

Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995 Jun;38(6):727-35.

Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Ann Rheum Dis 2011;70:404-13.

Genant HK, Jergas M, Palermo L, Nevitt M, Valentin RS, Black D, et al. Comparison of semiquantitative visual and quantitative morphometric assessment of prevalent and incident vertebral fractures in osteoporosis The Study of Osteoporotic Fractures Research Group. J Bone Miner Res 1996;11:984-96.

Guyatt GH, Norris SL, Schulman S, Hirsh J, Eckman MH, Akl EA, et al. American College of Chest Physicians. Methodology for the development of antithrombotic therapy and prevention of thrombosis guidelines: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. Chest. 2012 Feb;141(2 Suppl):53S-70S.

Hox JJ. Multilevel analysis. Techniques and applications. New York: Routledge; 2010.

Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models.J Stat Softw 2017;82:1-26. https://doi.org/10.18637/jss.v082.i13

Little RJA, Rubin DB. Statistical analysis with missing data (2nd ed.). Hoboken: Wiley; 2002.

R Core Team. R: A language and environment for statistical computing. Vienna; R Foundation for Statistical Computing; 2020. https://www.R-project.org/

Ritchie DM, Boyle JA, McInnes JM et al. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. Q J Med 1968; 37: 393-406.

Rubin DB. Causal Inference Using Potential Outcomes. J Am Stat Assoc 2005;100:322-31. http://doi.org/10.1198/016214504000001880

Sharp JT, Young DY, Bluhm GB et al. How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis? Arthritis Rheum 1985; 28: 1326-35.

Snijders TAB, Bosker RJ. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. New York: SAGE; 2011.

Van Buuren S. Flexible Imputation of Missing Data. Boca Raton: CRC Press; 2012.

Van der Heijde DM, Van Riel P, Nuver-Zwart IH, Gribnau FW, Van der Putte LB. Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. Lancet 1989; 1: 1036-8.

Van Gestel AM, Prevoo ML, Van 't Hof MA, Van Rijswijk MH, Van de Putte LB, Van Riel PL. Development and validation ofthe European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum. 1996 Jan;39(1):34-40.

Van Gestel AM, Anderson JJ, Van Riel PL, Boers M, Haagsma CJ, Rich B, Wells G, Lange ML, Felson DT. ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. American College of Rheumatology European League of Associations for Rheumatology. J Rheumatol. 1999 Mar;26(3):705-11.

Wells, G.A., Boers, M., Shea, B., Brooks, P.M., Simon, L.S., Strand, C.V., et al. (2005). Minimal disease activity for rheumatoid arthritis: a preliminary definition. J Rheumatol, 32, 2016-24.

10 Appendix A: secondary outcome measures that need further specification

Remission and minimal disease activity (MDA), according to ACR/EULAR criteria
Remission is defined according to the Boolean Based definition of remission in clinical trials. A patient must satisfy a tender joint count ≤1, a swollen joint count ≤1, CRP ≤1 mg/dL, and a patient global assessment ≤1 (on a 0-10 scale). MDA is defined as a DAS28 <2.6 (Felson et al., 2011; Wells et al., 2005).

ACR response
This index is described in the protocol. To calculate ACR50 and 70, the '20%' criterion is replaced by 50 resp 70%.

Vital signs
Systolic and diastolic blood pressure, pulse rate, weight and height.

Physical examinations
Physical examinations were performed at the discretion of the treating physician. The results will be summarized in a table per treatment group. All abnormalities will be reported.

Bone mass assessment
Bone mass of lumbar spine and total hip is assessed by Dual-energy X-ray Absorptiometry (DEXA), and expressed as g/cm^2 for analysis.
In addition, description of bone mass at femoral neck, and bone mass of all sites as T-score.

Vertebral fracture analysis (by DEXA, or alternatively, by assessment of lateral spine X-rays)
Vertebrae are viewed on transversal image and height loss (%) is scored according to Genant:
Grade 0: <20%; Grade 1: 20-25%; Grade 2 ≥25-40%; Grade 3 >40%.
baseline: prevalent fracture: number of vertebrae with at least Grade 1;
description of distribution of % height loss in vertebrae of patients with a prevalent fracture.
follow up: incident fracture: sum of:
a. number of vertebrae with baseline Grade 0 increasing to Grade 1-3
(% height loss <20% increasing to >20%)
b. number of vertebrae with baseline Grade >0 increasing in Grade
(Grade 1 increasing to 2 or 3; Grade 2 increasing to 3)

Medication adherence

Medication adherence according to pill count is calculated as follows: if the number of capsules dispensed is D, the treatment period in number of days is P, and the number of capsules returned is R, medication adherence (expressed as %) is calculated as: 100*(D-R)/P. In alignment with the literature, good medication adherence is defined as an intake of at least 80% of the prescribed doses.

For medication adherence according to electronic caps the assumption is made that a bottle opening is equal to the intake of one capsule. If the number of days that the cap wis not opened is O and the treatment period in number of days is P, adherence according to electronic caps (expressed as %) is calculated as: 100*(P-O)/P.

11 Appendix B – Modifications from the protocol

Definition of AESI

We refined the definitions of AESI in the protocol because it was not always completely clear when we were adjudicating the blinded events. The protocol text has been updated in order to make binding decisions on whether an event is AESI or not. The following changes and clarifications to the definition of an AESI were made:

- 'Stop of antirheumatic treatment for AE' was one of the options to define an event as an adverse event of special interest. However, the event 'stop of antirheumatic treatment for AE' is also taken into account as potential harm confounder in the analyses. We deleted this event from the definition of AE of special interest because we don't want to double count this event. So, the event ''any AE (except loss of efficacy, worsening of disease) that leads to the definite cessation of one of the antirheumatic drugs, including trial medication'' was changed to ''any AE (except loss of efficacy, worsening of disease) that leads to the definite cessation of trial medication''.
- 'Infection requiring antibiotic treatment' was discussed because the definition is not clear about topical treatments, and treatments for fungi and viruses. Therefore, we decided that:
    - For the purpose of classification, 'antibiotic' includes any specific treatment aimed at microorganisms to treat an infection: so this includes topical, antiviral, antifungal, antiparasitic treatment.
    - (Surgical) procedures with concomitant antibiotics: not (initially) an infection, so not AESI; AESI only when the procedure was done for a primary infection (e.g. abscess). Surgical complications treated with antibiotics are also not coded as AESI, because it is usually impossible to determine whether the antibiotic was given as prophylaxis or as treatment.
      Note: any surgery is of necessity (hospital admission) already an SAE.
- A cardiovascular event (myocardial infarction, cerebrovascular event, peripheral arterial vascular event);
    - Anything falling outside this definition is not an AESI.
      So venous hemorrhoids, venous ulcus cruris, thrombosis, pulmonary embolism, subdural hematoma, and traumatic hemorrhagic brain lesions are not coded as AESI. Also cardiac rhythm disturbances, valve lesions, cardiac insufficiency, cardiomyopathy, and encephalopathy, are not coded as AESI unless part of an acute ischemic episode.
- Newly occurring hypertension requiring drug treatment;
- Newly occurring diabetes mellitus requiring drug treatment;
- Symptomatic bone fracture requiring treatment;
    - AESI is coded as 'no' if the vertebral fracture was asymptomatic, and 'yes' for all nonvertebral fractures, unless expressly indicated that the fracture was asymptomatic.
      Note that vertebral fracture is also captured as secondary outcome through DEXA or spine radiographs.

- Newly occurring cataract or glaucoma.
    - Cataract and glaucoma are assumed to be bilateral.
      A medical history of the disease in one eye counts for both, so no AESI during the trial for procedures on the contralateral eye. We count bilateral procedures for new disease as one (not two) AESI.
- ''Newly occurring diabetes or hypertension'': no problems in adjudication.

10.2 Secondary endpoints

The measurements severity and duration of fatigue due to RA are included in the initial protocol, but not in the assessment. Severity of fatigue is part of the RA Impact of Disease (RAID) questionnaire and will be reported as described in section 1.4.2. Duration of fatigue was not included in the assessment and will not be reported.

The number of patients retained on study drug over time was not described as secondary endpoint in the protocol. However, this proportion is measured and will be presented with survival analysis.

12 Appendix C. Report of blinded data review meetings

Blinded Review of deviations leading to exclusion
from the Per-Protocol Analysis

Report of Review Meetings 9 June, 14 July, 17 November 2020 and 8 February 2021.
Final version dd 2 March 2021.

Attendees: Maarten Boers (PI/MM), Linda Hartman (PhD candidate), Hans Bijlsma (SAC member), Nick van der Bulk (CR2O, Head Clin Ops), Liza Bakker (sr. CRA), Linda Doerwald (Linical, Program Data manager), Leonie Middelink (Ops lead Gloria)

Documents used

- Clinical study protocol version 4.0 12 February 2019
- Medical monitor plan dd. 05 April 2018
- PD tacker dd. 18 February 2021
- Manual listings 11 February 2021

1. Investigational Medicinal Product (IMP) deviations

The list of patients missing >20% of capsules per period (up to 3 months and up to 24 months) was reviewed. It was agreed that for the Per Protocol (PP) population adherence of at least 80% up to 3 months is required.
Review list :

1. Adherence should be at least 80% first 3 months done (between baseline and V2), if adherence is < 80%, the patients are excluded from the PP group. For the complete overview, see attachment PP vs ITT overview.
2. The periods of non-compliance after 3 months up to 24 months are left out of scope as they do not influence the PP population that applies only to the first 3 months.

2. Protocol Deviations

The PD tracker (dd 04 Feb 2021, later updated with version 18 February 2021) was reviewed, and following deviations were discussed:

Possible unblinding

Patients 107005 and -021 opened the capsules themselves.
As the unblinding took place in the first 3 months, they are out of the PP group.
Patient 107005 ICF signed on 31 Jan 2017, unblinding 24 Mar 2017; 107021 ICF signed on 22 aug 2018 and unblinding 6 Nov 2018

Prednisone open label use

The PD tracker was reconciled with the Conmeds page in the manual listings. This has been done after final coding. This list has been crossed checked with the PD tracker, and for all prohibited prednisolone open label use an PD form has been created. If this prohibited open label use took place in the 30 days before baseline, or between ICF signature date and V2, the patient is excluded

from the PP population (see prohibited meds sections). Other prednisolone open label use (after V2) will be flagged for analysis, but does not affect the PP selection.

In the coded conmeds file, the H02 ATC codes were selected. And the following rules were applied:

- IV GCs are always a Protocol Deviation (if before stop of IMP).
- Oral GCs are allowed for a maximum period of 3 weeks, for max 4 times, during IMP use.
- Intra-Articular and Intra Bursal are allowed a maximum of 4 times during IMP use
- Intra Muscular GCs are allowed to a maximum of 2 times during IMP use.

IMP deviations
If a patient misses a substantial percentage of the doses in a certain period (e.g. patient 116004 and 501009) after month 3, this does not affect the PP population, see above.

Time window deviations
As the time window deviations all have taken place after V2, this does not affect the PP evaluation.

Missed assessments
Only missing DAS-scores during the first 3 months of the trial will lead to exclusion from PP.
A DAS28-CRP can be used instead of the DAS28ESR. It was decided not to impute the DAS-ESR with a RAPID3. For the following visits the DASESR will be imputed with a DASCRP.

PatiëntID	Imputation	Visite
104016	DAS28ESR imputed with DAS28CRP	2
106004	DAS28ESR imputed with DAS28CRP	2
110002	DAS28ESR imputed with DAS28CRP	1
110004	DAS28ESR imputed with DAS28CRP	2
113007	DAS28ESR imputed with DAS28CRP	2
113010	DAS28ESR imputed with DAS28CRP	1
114013	DAS28ESR imputed with DAS28CRP	2
117006	DAS28ESR imputed with DAS28CRP	2

Process deviations
Deviations such as SAE timelines, ICF signatures etc. do not lead to exclusion.

Eligibility criteria
All the deviations regarding eligibility criteria have been reviewed. Patients with actual deviations regarding violation of selection criteria are excluded from the PP. Process deviations relating to e.g. the Informed Consent process will not.

3. Other RA medication

The type of medication, timing and exposure to the RA medication next to the IMP has been reviewed based on the manual listings.

The objective of the PP analysis in the first three months is to study the effects of low-dose prednisolone in a population without co-interventions. As the impact of concurrent RA med use is so substantial, patients who start or change prohibited meds (biologicals, DMARDs, GCs) between ICF and V2 of the trial are excluded from the PP population as well. As per protocol, start with other RA meds was not absolutely prohibited, we changed the term “per protocol population”, “per protocol population, also excluding initial changes in antirheumatic medication”. This terminology has also been updated in the final SAP.

4. Other reviews to consider

- Safety reporting: cases of increase in disease activity (DA) were reported as AEs. No increase in DA reported as AE (based in coding file 5feb2021). Discussion pat 104008, AE# 9, swollen ancle left and right (NB site 114 3x painful joints / knee). All these AEs have been queried and the site confirmed it was an actual separate AE.

- The check of efficacy data for patients with an early end of treatment has been incorporated in the SAP.

The wording of the definitions of AEs of SI has been finetuned. Although this does not affect the PP evaluation, this has been discussed (as well as with the SAC), and the new rules for stop of RA meds and “anti-biotics” are to be applied after data export and before analysis. This has also been added to the SAP in the section “deviations from protocol”.

The final list of patients included in the PP population is shown in Appendix D, below.

13 Appendix D. Gloria final PP population, also excluding initial changes in antirheumatic medication

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
101001	In				Out	Out	Out
101002	In						
101003	In				Out	Out	Out
101004	In						
101005	Out			DAS28 v2		Out	Out
101006	In			DAS28 v2	Out		Out
101007	In						
101008	In			DAS28 V2			Out
101009	In						
101010	In						
101011	Out						Out
101012	Out			DAS28 v2		Out	Out
101013	In					Out	Out
101014	In						
101015	In						
101016	In						
101017	In			DAS28 v1		Out	Out
101018	Out			DAS28 v2	Out	Out	Out
101019	In					Out	Out
101020	In					Out	Out
101021	In						
101022	In						
101023	In						
101024	In						
101025	In					Out	Out
101026	In						
101027	In						
101028	Out						Out
101029	In						
101030	In					Out	Out
101031	In						
101032	In						
101033	In						
101034	In						
101035	In						
101036	In						
101037	In						
101038	In					Out	Out
101039	In						
102001	In					Out	Out
102002	In						

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
102003	In						
102004	In				Out	Out	Out
102005	Out						Out
102006	In						
102007	In						
102008	In						
102009	In						
104001	In					Out	Out
104004	In					Out	Out
104005	In						
104006	In						
104007	In					Out	Out
104008	In						
104009	In						
104010	In						
104011	In					Out	Out
104012	In						
104013	In						
104014	In					Out	Out
104015	In					Out	Out
104016	In					Out	Out
104017	In						
104018	In						
104019	In						
106001	In					Out	Out
106002	In						
106003	In						
106004	In						
106005	In						
106006	In						
106007	In					Out	Out
106008	In						
107001	In						
107002	In		GC use first 3 months			Out	Out
107003	In						
107004	In						
107005	In	Unblinding < m3		DAS28 v2			Out
107006	In						
107007	Out			DAS28 v2			Out
107008	In						
107009	In			DAS28 v2		Out	Out
107010	In						
107011	In					Out	Out
107012	In						
107013	In					Out	Out

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
107014	In					Out	Out
107015	In					Out	Out
107016	In						
107018	In						
107019	In					Out	Out
107020	In					Out	Out
107021	In	Unblinding < m3					Out
107022	In						
107023	In					Out	Out
107024	In					Out	Out
108001	In						
108002	In						
108003	In					Out	Out
108004	In						
108005	In						
108006	Out			DAS28 v2		Out	Out
108007	In						
108008	In						
108009	In						
108010	In						
108011	In						
108012	Out						Out
108013	In						
108014	In						
108015	Out			DAS28 v2			Out
108016	In						
108017	In						
108018	In						
108019	Out			DAS28 v2		Out	Out
108020	Out					Out	Out
108021	In					Out	Out
108022	In						
108023	In					Out	Out
108024	In						
108025	Out						Out
108026	In					Out	Out
108027	Out			DAS28 v2		Out	Out
108028	In						
108029	In						
108030	In						
108031	In					Out	Out
108032	In						
108033	In			DAS28 v2		Out	Out
108034	In					Out	Out
108035	Out					Out	Out
108036	In						

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
108037	In						
108038	In					Out	Out
108039	Out			DAS28 v2		Out	Out
108040	In					Out	Out
108041	In						
108042	Out						Out
108043	In					Out	Out
109001	In						
109002	In					Out	Out
109003	In						
109004	In						
109005	In						
109007	In					Out	Out
109008	In						
109009	In						
109010	In						
109011	In						
109012	In					Out	Out
109013	In					Out	Out
109014	In						
109015	In					Out	Out
109016	Out						Out
109017	In						
109018	In						
109019	In					Out	Out
110001	In						
110002	In						
110003	In						
110004	In						
110005	In						
110006	Out			DAS28 v2			Out
110007	Out						Out
110008	In					Out	Out
110009	In						
110010	Out						Out
111001	In						
111002	In						
111003	In						
111004	In						
111005	In						
111006	In						
111007	In						
111008	In						
111009	In						
112001	In						
112002	In						
112003	In						

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
112004	In						
112005	In					Out	Out
112006	In			DAS28 V2			Out
112007	In						
112008	Out						Out
112009	In						
112010	In						
112011	In						
112012	In						
112013	In						
112014	In						
112015	In					Out	Out
112016	In					Out	Out
112017	In						
112018	In						
112019	In					Out	Out
112020	In						
112021	In					Out	Out
112022	In						
112023	In						
112024	In						
112025	In						
112026	In						
112027	In						
112028	In						
112029	In						
112030	In					Out	Out
112031	In						
112032	In						
112033	In		Pred < 30d baseline				Out
112034	In						
112035	In						
112036	In						
112037	Out						Out
112038	Out						Out
112039	In		Pred < 30d baseline			Out	Out
112040	In						
112041	In						
112042	In					Out	Out
112043	Out			DAS28 v2			Out
112044	In						
112045	In			DAS28V2			Out
112046	In						
112047	In						
112048	Out						Out
112049	In						

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
112050	In					Out	Out
112051	In						
112052	In						
112053	In						
112054	In					Out	Out
113001	In						
113002	Out			DAS28 v2			Out
113003	In						
113004	In						
113005	In					Out	Out
113006	In					Out	Out
113007	In						
113008	In						
113009	In						
113010	In					Out	Out
113011	In						
113012	In			DAS28 v2		Out	Out
114001	In			DAS28 v1		Out	Out
114002	In						
114003	In						
114004	In					Out	Out
114005	Out						Out
114006	Out						Out
114007	In						
114008	In						
114009	In						
114010	In					Out	Out
114011	In						
114012	In					Out	Out
114013	In						
114015	In			DAS28V2	Out		Out
114016	In						
115001	In						
115002	In						
115003	In						
116001	In						
116002	In						
116003	In						
116004	In				Out		Out
116005	In						
116006	In						
116007	In						
116008	In						
116009	In						
116010	In						
116011	Out			DAS28 v2			Out
116012	In						

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
116013	In			DAS28 v2			Out
116014	In						
116015	In						
117001	In						
117002	In						
117003	In					Out	Out
117004	In						
117005	In						
117006	In						
117007	In						
117008	In						
117009	In						
117010	In						
201001	In						
201002	In						
204001	Out			DAS28 v2			Out
204002	In						
204003	In						
204004	In						
204005	In						
204006	In						
204007	In						
205001	Out					Out	Out
301001	In						
301002	In						
301003	In						
301004	In						
301005	In						
301006	Out			DAS28 v2			Out
301007	In						
301008	In						
301009	In			DAS28 V2			Out
302001	In						
303001	In						
304001	In						
304002	Out						Out
304003	In						
304004	In						
304005	In						
304006	In						
304007	Out						Out
304008	In						
305001	In						
305002	In						
305003	In						
305004	In						
305005	In						

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
305006	In					Out	Out
305007	In						
401001	In						
401002	In						
501001	In						
501002	In						
501003	In						
501004	In						
501005	In						
501006	In						
501007	In						
501008	In						
501009	In						
501010	In						
501011	In						
501012	In						
501013	In						
501014	In						
501015	In						
501016	In						
501017	In						
501018	In						
501019	In						
501020	In						
501021	In						
501022	In						
501023	In						
501024	In					Out	Out
501025	In						
501026	In						
501027	In						
501028	In					Out	Out
501029	In						
501030	In					Out	Out
501031	In						
501032	In						
501033	In						
501034	In						
501035	In					Out	Out
501036	In						
501037	In						
501038	In						
501039	In						
501040	In						
501041	In						
501042	In						
501043	In					Out	Out

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
501044	In						
501045	In						
501046	In						
501047	In						
501048	Out			DAS28 v2			Out
501049	In						
501050	In						
501051	Out					Out	Out
501052	Out			DAS28 v2			Out
501053	In						
501054	In						
501055	Out						Out
501056	In						
501057	In						
501058	In						
501059	In						
501060	In						
601001	Out			DAS28 v2			Out
601002	Out			DAS28 v2			Out
601003	In						
601004	In						
601005	Out			DAS28 v2		Out	Out
601006	In						
601007	In						
601008	In						
601009	Out			DAS28 v2			Out
601010	In						
601011	Out			DAS28 v2			Out
601012	In						
701001	In						
701002	In						
701003	In						
701004	In						
701005	In						
701006	Out			DAS28 v2			Out
701007	In						
701008	In						
701009	In						
701010	Out						Out
701011	In						
701012	In					Out	Out
701013	In						
701014	In						
701015	In						
701016	In						
701017	In						
701018	In						

Patient	IMP <3 m	Possible unblinding	Prohib pred use	Missed Ass	Eligibility Vio	Other RA meds ICF-V2	Final
701019	Out			DAS28 v2			Out
701020	In						
701021	Out						Out
701022	In						
701023	Out						Out
701024	Out			DAS28 v2		Out	Out
701025	In						
701026	In					Out	Out
701027	In						
701028	In						
701029	In						
701030	Out			DAS28 v2			Out
701031	In						
701032	In						
701033	Out			DAS28 v2			Out
701034	In						
701035	In					Out	Out
701036	In						
701037	In						
701038	In						
701039	Out			DAS28 v2			Out
701040	In						
701041	In						
701042	In						
701043	In						
701044	Out					Out	Out
701045	In					Out	Out
702001	Out						Out
702002	In						
702003	Out			DAS28 v2			Out
702004	Out						Out
702005	In						
702006	In						
702007	Out						Out
702008	In						
702009	In						
702010	Out						Out
702011	Out						Out
total	**60**	**2**	**3**	**42**	**7**	**90**	**147**