Statistical Analysis Plan

Protocol number:	Sobi.ANAKIN-401
NCT Number:	03002974
Title:	A randomized, double-blind, active-control, multicenter, efficacy and safety study of 2 dose levels of subcutaneous anakinra compared to intramuscular triamcinolone in the treatment of acute gouty arthritis, followed by an extension period of up to 2 years
SAP version:	Final 1.0
Date:	17 August 2018

Table of contents

1 List of abbreviations and definition of terms 6
2 Introduction 8
3 Study objectives and endpoints 8
3.1 Primary objective 8
3.2 Secondary objectives 8
3.2.1 Exploratory objectives 9
3.3 Study endpoints 9
3.3.1 Primary efficacy endpoint and secondary endpoints supporting the primary objective 9
3.3.1.1 Primary efficacy endpoint 9
3.3.1.2 Secondary endpoints supporting the primary objective 9
3.3.2 Secondary efficacy endpoints 10
3.3.3 Safety endpoints 10
3.3.4 Serum concentrations 11
3.3.5 Exploratory endpoints 11
4 Study methods 11
4.1 Overall study design and plan 11
4.2 Selection of study population 13
4.3 Method of treatment assignment and randomization 14
5 Changes from protocol affecting the statistical analysis 15
6 Sequence of planned analysis 15
6.1 Main efficacy analysis 15
6.2 Final analysis 16
6.3 Reporting 16
7 Sample size determination 16
8 Analysis populations 17
8.1 Intention-to-treat (ITT) 17
8.2 Per-protocol population (PP) 17
8.3 Safety population 17
9 General considerations for statistical analysis 18
9.1 Handling of missing data and outliers 18
9.2 Multicenter studies 19
9.3 Multiple comparisons and multiplicity 19
10 Patient disposition 19

11 Demographics and baseline characteristics 20
11.1 Demographics 20
11.2 Baseline characteristics 20
11.3 Medical and surgical history 20
12 Prior and concomitant medication 21
13 Study drug compliance 21
14 Efficacy analyses 22
14.1 Primary efficacy endpoint 22
14.1.1 Subgroup analyses 23
14.1.2 Sensitivity analysis of the primary endpoint 23
14.1.2.1 Sensitivity to subject population 24
14.1.2.2 Sensitivity to missing data handling 24
14.1.2.3 Tipping point analysis 24
14.2 Secondary endpoints supporting the primary objective 26
14.2.1 Patient-assessed pain intensity in the index joint on a Visual analogue scale (VAS) 26
14.2.2 Patient-assessed pain intensity in the index joint on a 5-point Likert scale 27
14.3 Secondary efficacy endpoints 28
14.3.1.1 Pain intensity for anakinra dose groups, 100 and 200 mg 28
14.3.1.2 Onset of effect 28
14.3.1.3 Response 29
14.3.1.4 Resolution of pain 29
14.3.1.5 Rescue medication 30
14.3.1.6 Physician’s assessment of global response to treatment 31
14.3.1.7 Physician’s assessment of clinical signs 31
14.3.1.8 Patient´s assessment of global response to treatment 30
14.3.1.9 Inflammatory biomarkers 32
14.3.1.10 Subsequent flares 32
15 Serum concentrations of Interleukin 1 receptor antagonist (IL-1Ra)/anakinra 34
16 Safety analyses 34
16.1 Drug exposure 34
16.2 Adverse events 34
16.2.1 Serious adverse events 35
16.2.2 Adverse events leading to withdrawal 35
16.2.3 Deaths 36
16.3 Laboratory data 36

16.4 Vital signs ... 36
16.5 Electrocardiogram ... 36
16.6 Anti-drug antibodies ... 36
16.6.1 Presence of anti-drug antibodies ... 37
16.6.2 Anti-drug antibodies versus safety ... 37
16.6.3 Anti-drug antibodies versus efficacy ... 37
16.6.4 Anti-drug antibodies versus serum concentration of IL-1Ra/anakinra ... 38
17 Exploratory endpoints ... 38
17.1 Short Form (36) Health Survey (SF-36) ... 38
17.2 EuroQol 5 dimensions 5 levels (EQ-5D-5L) ... 38
17.3 Work productivity and activity impairment:specific health problems (WPAI:SHP) ... 39
17.4 Health care resource utilization ... 39
17.5 Exploratory inflammatory biomarkers ... 39
18 Derived and computed variables ... 39
18.1 Unpermitted rescue medication ... 39
18.2 Age ... 40
18.3 Body mass index (BMI) ... 40
18.4 Age at diagnosis of gouty arthritis ... 40
18.5 Disease duration ... 40
18.6 Start of medical and surgical history ... 41
18.7 eGFR category ... 41
18.8 Obesity ... 41
18.9 Cardiovascular disease ... 41
18.10 Hyperlipidemia ... 41
18.11 Cerebrovascular disease ... 42
18.12 Start of prior and concomitant medication ... 42
18.13 Glucocorticoid use ... 42
18.14 ULT use ... 42
18.15 Study drug compliance ... 42
18.16 Time windows for pain intensity assessments (Visual analogue scale – VAS- and Likert) ... 43
18.17 Pain intensity on Visual analogue scale (VAS) ... 44
18.18 Pain intensity on Likert ... 44
18.19 Onset of effect ... 44
18.20 Response ... 45
18.21 Resolution of pain ... 45

18.22 Time to rescue medication 45
18.23 Drug exposure 46
18.24 Adverse event start/stop day 47
18.25 Treatment emergent abnormal vital signs 48
18.26 Immunogenicity 48
18.27 Short Form (36) Health Survey (SF-36) 48
18.28 EuroQol 5 dimensions 5 levels (EQ-5D-5L) 49
18.29 Work productivity and activity impairment:specific health problems (WPAI:SHP) 50
18.30 Health care resource utilization 51
19 References 52

Table of tables

Table 1 eGFR categories 41
Table 2 Time windows for pain intensity assessments (VAS and Likert) 43

Table of figures

Figure 1 Study design 12
Figure 2 SF-36 domain scores ..... **Error! Bookmark not defined.**

1 List of abbreviations and definition of terms

Abbreviation	Definition
ACR	American College of Rheumatology
ADA	Anti-drug antibodies
AE	Adverse event
ATC	Anatomical Therapeutic Chemical Classification System
BMI	Body mass index
CRF	Case report form
CRO	Contract research organization
CRP	C-reactive protein
CSR	Clinical study report
ECG	Electrocardiogram
eCRF	Electronic case report form
eDiary	Electronic diary
eGFR	Estimated glomerular filtration rate
ePRO	Electronic PRO
EQ-5D-5L	EuroQol 5 dimensions 5 levels
EULAR	European League Against Rheumatism
FDA	Food and Drug Administration
GFR	Glomerular filtration rate
HR	Heart rate
HRQL	Health related quality of life
ICF	Informed consent form
ICH	International Council of Harmonization
IL-1Ra	Endogenous interleukin 1 receptor antagonist
IL-6	Interleukin-6
IL-8	Interleukin-8
i.m.	Intramuscular
IMP	Investigational medicinal product
ITT	Intention-to-treat
IWRS	Interactive web response system
LS	Least square
LSM	Least square means
MAR	Missing at random
MCS	Mental component summary
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed model repeated measures

Abbreviation	Definition
MNAR	Missing not at random
MPO	Myeloperoxidase
NAb	Neutralizing anti-drug antibodies
No.	Number
NSAID	Non-steroidal anti-inflammatory drugs
PCS	Physical component summary
PP	Per-protocol
PRO	Patient reported outcome
PT	Preferred term
SAA	Serum amyloid A
SAE	Serious adverse event
SAP	Statistical analysis plan
s.c.	Subcutaneous
SD	Standard deviation
SDTM	Study data tabulation model
SF-36	Short Form (36) Health Survey (acute, Version 2)
Sobi	Swedish Orphan Biovitrum
SOC	System organ class
SOP	Standard operating procedure
TEAE	Treatment-emergent adverse event
TIA	Transient ischemic attack
ULT	Urate lowering therapy
VAS	Visual analogue scale
WPAI:SHP	Work productivity and activity impairment:specific health problems

2 Introduction

This SAP describes the planned analyses and reporting for the Sobi study Sobi.ANAKIN-401, A randomized, double-blind, active-control, multicenter, efficacy and safety study of 2 dose levels of subcutaneous anakinra compared to intramuscular triamcinolone in the treatment of acute gouty arthritis, followed by an extension period of up to 2 years.

This phase II study is being completed to assess the efficacy and safety of anakinra for the treatment of acute gouty arthritis in patients for whom NSAIDs and colchicine are either contraindicated or not appropriate due to anticipated changes in patient status, not tolerated or do not provide an adequate response.

The purpose of this SAP is to outline the planned analyses to be completed to support the CSR for the study Sobi.ANAKIN-401. The planned analyses identified in this SAP will be included in regulatory submissions and/or future manuscripts. Also, exploratory analyses not necessarily identified in this SAP may be performed to support the clinical development program. Any post-hoc or unplanned analyses not identified in this SAP will be clearly identified in the respective CSR.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the FDA and International Council of Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH): Guidance on Statistical Principles in Clinical Trials. All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association and the Royal Statistical Society, for statistical practice.

3 Study objectives and endpoints

3.1 Primary objective

The primary objective of the study is to evaluate the efficacy of anakinra compared to triamcinolone acetonide 40 mg i.m. injection, hereafter referred to as triamcinolone, with respect to patient-assessed pain intensity in the treatment of a gouty arthritis flare (based on patients' first flare treated in the study).

3.2 Secondary objectives

The secondary objectives of the study are:

- To evaluate the primary endpoint and the secondary endpoints supporting primary objective for the 2 different anakinra dose groups compared to triamcinolone in the treatment of the first gouty arthritis flare.
- To evaluate the time to onset of effect, time to response, time to pain resolution, time to rescue medication, physician's assessment of global response and clinical signs, patient´s

assessment of global response, inflammatory biomarkers, and safety of anakinra compared to triamcinolone in the treatment of the first gouty arthritis flare both for the combined anakinra group (100 and 200 mg) and the 2 different anakinra dose groups.
- To evaluate the primary and secondary endpoints and safety of the combined anakinra group (100 and 200 mg), the 2 different anakinra dose groups and triamcinolone in the treatment of subsequent flares. Subsequent flares are flares occurring after the first flare in the study and within 52 weeks of randomization of the last randomized patient in the study. However the extension will be a maximum of 2 years (104 weeks) after randomization for the individual patient in the study.

3.2.1 Exploratory objectives

The exploratory objectives of the study are:

- To evaluate the effect of anakinra compared to triamcinolone on HRQL and health care resource utilization in the treatment of the first gouty arthritis flare.
- To evaluate the effect of anakinra compared to triamcinolone on exploratory inflammatory biomarkers in serum in the treatment of gouty arthritis flares. (Statistical analyses of exploratory inflammatory biomarkers of more exploratory nature will not be described in this SAP and these results may be reported separately).
- To analyze for genetic factors potentially contributing to the patient's response to anakinra, safety and tolerability. Such genetic factors may include genes within the target pathway, or other genes believed to be related to the response to anakinra. The statistical analysis of genetic factors will not be described in this SAP and these results may be reported separately.

3.3 Study endpoints

3.3.1 Primary efficacy endpoint and secondary endpoints supporting the primary objective

3.3.1.1 Primary efficacy endpoint

The primary efficacy endpoint is:

- The change in patient-assessed pain intensity from baseline to 24 to 72 hours (average of the assessments performed at 24, 48, and 72 hours) for the first flare.

3.3.1.2 Secondary endpoints supporting the primary objective

The following secondary endpoints will be evaluated to support the primary objective:

- Change from baseline in patient-assessed pain intensity in the index joint as measured by VAS at 6, 12, 18, 24, 36, 48, 72 hours, and Days 5, 6, 7, and 8 for the first flare.

- Change from baseline in patient-assessed pain intensity in the index joint as measured by a 5-point Likert scale at 6, 12, 18, 24, 36, 48, 72 hours, and Days 5, 6, 7, and 8 for the first flare.

3.3.2 Secondary efficacy endpoints

The following secondary endpoints will be assessed for the first flare and subsequent flares. All comparisons versus baseline (pre-dose measurement at Visit 1) will be made against the baseline of the first and subsequent flares, respectively.

- Time to onset of effect (defined as ≥20 % change from baseline pain intensity on VAS).
- Time to response (defined as ≥50 % change from baseline pain intensity on VAS).
- Response (defined as ≥50 % change from baseline pain intensity on VAS) at 24, 48 and 72 hours, Day 8 and Day 15 (yes/no).
- Resolution of pain (defined as <10 mm on VAS) at 72 hours, Day 8 and Day 15 (Yes/No).
- Time to resolution of pain (defined as <10 mm on VAS).
- Time to intake of rescue medication from first IMP administration.
- Type and number of occasions of intake of rescue medication taken from first IMP administration to Day 8.
- Patient´s assessment of global response to treatment (5-point Likert scale) at 72 hours, Day 8 and Day 15.
- Physician's assessment of global response to treatment (5-point Likert scale) at 72 hours, Day 8 and Day 15.
- Physician's assessment of clinical signs at 72 hours, Day 8 and Day 15 in index joint tenderness, swelling and erythema.
- Change from baseline in the inflammatory biomarkers CRP and SAA at 72 hours, Day 8 and Day 15.

Additional assessment for the subsequent flares:

- Change in patient-assessed pain intensity in the index joint as measured by VAS score from baseline to 24 to 72 hours (average of the assessments performed at 24, 48 and 72 hours).
- Change from baseline in patient-assessed pain intensity in the index joint as measured by VAS score at 6, 12, 18, 24, 36, 48, 72 hours and Days 5, 6, 7 and 8.
- Change from baseline in patient-assessed pain intensity in the index joint as measured by a 5-point Likert scale at 6, 12, 18, 24, 36, 48, 72 hours and Days 5, 6, 7 and 8.

3.3.3 Safety endpoints

The study has the following safety endpoints:

- Occurrence of AEs.
- Laboratory safety assessments.

- Vital signs .
- Occurrence of ADA against anakinra at Visit 1 (before IMP administration) and at Day 8, Day 15, Day 28 and Week 12 (the assessment at Day 28 and Week 12 will only be performed if no subsequent flare has occurred at that timepoint).

3.3.4 Serum concentrations

The following endpoint will be assessed to evaluate anakinra serum concentrations:

- Serum concentration of IL-1Ra/anakinra at baseline, 72 hours (just before IMP administration) and at Day 8, Day 15, Day 28 and Week 12 (the assessment at Day 28 and Week 12 will only be performed if no subsequent flare has occurred at that timepoint).

3.3.5 Exploratory endpoints

The following exploratory endpoints will be assessed for both the first and subsequent gouty arthritis flares. All comparisons versus baseline (pre-dose measurement at Visit 1) will be made against the baseline of the first and subsequent flares, respectively.

- Change from baseline in health related quality of life (SF-36® Health Survey, acute Version 2) at Day 8 and Day 15.
- Change from baseline in health related quality of life (EQ-5D-5L) at Day 8 and Day 15.
- Work productivity and activity impairment due to a gouty arthritis flare (WPAI:SHP Version 2.0) and health care resource utilization: number of days with hospitalization and number of unscheduled outpatient visits during the last week recorded at Day 8 and Day 15.
- Change from baseline in the exploratory inflammatory biomarkers IL-6, IL-8, calprotectin, neopterin and MPO in serum at 72 hours, Day 15 and Week 12.

4 Study methods

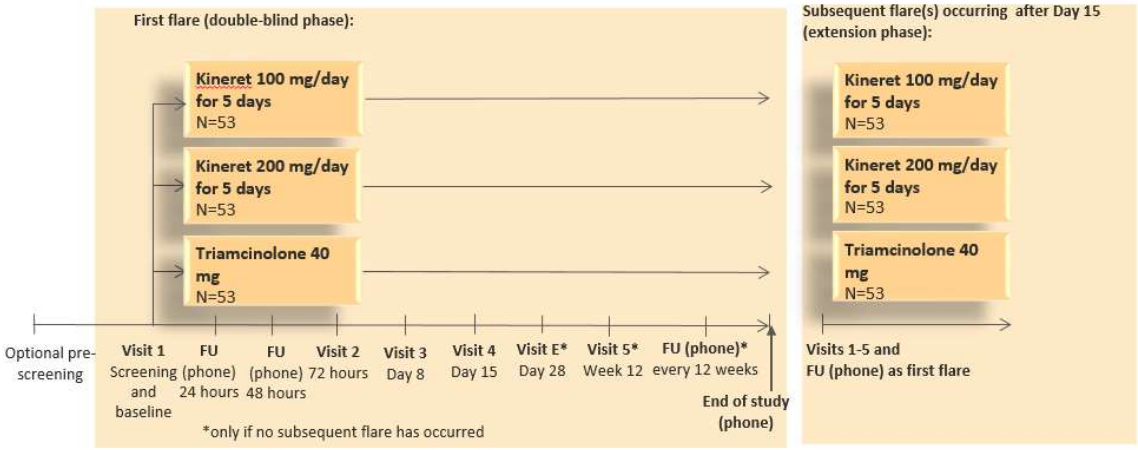
4.1 Overall study design and plan

The study consists of 3 periods: an optional pre-screening period, a double-blind treatment period and an extension period.

The treatment period of the first flare is double-blind, and the patients will be randomized to treatment with 100 mg anakinra, 200 mg anakinra or 40 mg triamcinolone acetonide in a 1:1:1 ratio. At Visit 1 of each flare, the triamcinolone or matching placebo will be given as one i.m. injection of 1 mL, and at each day of IMP administration period (Days 1 to 5), the anakinra and/or matching placebo will be given as 2 s.c. injections of 0.67 mL.

The double-blind period will be followed by an extension period. During this period, the patients will receive the same treatment for subsequent flares occurring after Day 15 of the latest flare. The extension period will continue for all patients until 52 weeks after randomization of the last randomized patient in the study. However the study treatment of an individual patient's flare will not start later than 2 years (104 weeks) after randomization of that patient. Thus the study will not exceed 104 + 12 weeks for any patient An overview of the study design is presented in Figure 1.

Figure 1 Study design



Abbreviations: FU, Follow-up; N, Number of patients.

At each baseline visit (first and subsequent flares), the joint that is most affected, i.e. most painful, will be defined as the index joint.

The assessment period for each flare will consist of 4 to 6 visits to the study center. The visits are the screening/baseline visit (Visit 1/Day 1), the 72-hour visit (Visit 2/Day 4), the 1-week visit (Visit 3/Day 8), the 2-week visit (Visit 4/Day 15), the 4-week visit (Visit E/Day 28) and the 3 month visit (Visit 5/Week 12). If a subsequent flare has been treated before the Day 28 visit, the Day 28 and Week 12 visits will be cancelled. If a subsequent flare has been treated before a Week 12 visit, that Week 12 visit will be cancelled. Subsequent flares occurring before the Day 15 visit of the latest flare (Visit 4) will be treated with permitted rescue medication, i.e. no IMP will be administered.

A follow-up by phone will be performed at 24 and 48 hours post first IMP administration to remind the patient to complete the assessments and recordings in the eDiary. A follow-up by phone will also be performed every 12 weeks following Visit 5 of the latest flare.

An end-of-study follow-up will be performed either;

1. 52 weeks after randomization of the last patient in the study

or

2. 2 years (104 weeks) after randomization of the individual patient

whichever occurs first.

If the IMP treatment started week 40 or later in case 1), or week 92 weeks in case 2), the Week 12 visit will replace the End-of-study follow-up by phone.

All patients will have their follow-up visits/phone calls, irrespective of treatment withdrawal or use of rescue medication.

Subsequent flares occurring before the Day 15 visit of the latest flare (Visit 4) will be treated with permitted rescue medication i.e., no IMP will be administered.

The AE reporting period for all AEs, including SAEs, begins upon receiving the first dose of IMP and ends at Visit E, 28 days after the first dose of IMP. SAEs will also be collected from signing of the ICF until the Week 12 visit (Visit 5). Furthermore, any SAE should be reported irrespective of the time of occurrence if a causal relationship between the event and the IMP(s) is suspected. If the patient is treated with IMP for a subsequent flare, the AE (including SAE) reporting period starts again when the first dose is administered.

Patients who have difficulty in tolerating their pain are allowed to take rescue medication after the 24-hour post-dose pain assessment but not within 6 hours before the 48- and 72-hour post-dose pain assessments. Permitted rescue medication includes topical ice/cold packs, paracetamol/acetaminophen (up to 1000 mg 4 times daily) and/or codeine (up to 50 mg 4 times daily) or short-acting tramadol (up to 100 mg 4 times daily).

Patients still having insufficient relief with the rescue medication listed above after the 72-hour post-dose pain assessment are allowed to take oral prednisone or prednisolone at a dose up to 0.5 mg/kg per day for a maximum of 5 days.

When all patients have conducted the Day 15 visit of the first flare (Visit 4), the database will be locked and the main efficacy analysis will be conducted. The main efficacy analysis includes primarily the statistical anlysis of all efficacy endpoints for the first flare. A second database lock will be conducted when all patients have completed the study.

4.2 Selection of study population

This study is designed to evaluate the efficacy and safety of anakinra in the treatment of gouty arthritis flares in adult male and female patients for whom NSAIDs and colchicine are either contraindicated or not appropriate due to anticipated changes in patient status, not tolerated or do not provide an adequate response. To ensure a reliable diagnosis of gouty arthritis, patients to be included in the study must meet the ACR/EULAR 2015 gout classification criteria (**Error! Reference source not found.**).

A history of ≥1 self-reported flare of gouty arthritis during the previous 12 months is required for inclusion in the study to ensure that the patients have an active gouty arthritis.

To ensure that a significant number of subsequent flares are treated in the study, the extension period will continue for all patients until 52 weeks after randomization of the last randomized

patient in the study or a maximum of 104 weeks after randomization of the individual patient in the study.

To be included, patients must have a patient-reported current ongoing flare of gouty arthritis characterized by baseline pain intensity in the index joint of ≥50 on a 0 to 100 VAS, as well as currently tender (≥1 on a 0-4-point Likert scale) and swollen (≥1 on a 0-4-point Likert scale). The onset of the current ongoing flare should be within 4 days prior to randomization.

Further, if the patient is on urate-lowering therapy, a stable dose and regimen for at least 2 weeks prior to randomization and expectance to remain on a stable dose and regimen for ≥2 weeks after administration of the first dose of study treatment is required.

Complete details of the patient inclusion and exclusion criteria are provided in CSP Section 7.3.1 and CSP Section 7.3.2.

Patients with data collection failure for primary endpoint are defined as patients who are randomized in the clinical study but have no values collected for patient-assessed pain intensity from baseline to 72 hours for the first flare. These patients will however continue in the study and be included in the ITT and safety populations. For each primary endpoint collection failures, one extra patient can be randomized in the study. Maximum 10 extra patients can be randomized under this protocol.

4.3 Method of treatment assignment and randomization

Up to totally 169 patients can be randomized in the study. Initially 159 patients will be randomized, 53 to triamcinolone and 106 to anakinra (53 to 100 mg and 53 to 200 mg). As described in Section 7, up to a maximum of 10 extra patients can be randomized. Patients will be randomized to treatment with anakinra 100 mg, anakinra 200 mg or triamcinolone 40 mg in a 1:1:1 ratio. The randomization will be stratified by ULT use (yes/no) and BMI (<30.0 or ≥30.0 kg/m^2) at inclusion.

For subsequent flares during the extension period, patients will continue to receive the same treatment with anakinra 100 mg, anakinra 200 mg or triamcinolone 40 mg.

An IWRS will be used for the randomization. Biostatistics at Perceptive will generate the randomization scheme for the IWRS, which will link sequential patient randomization numbers to treatment codes.

The randomization numbers will be generated in blocks. Within each block, equal numbers of patients will be allocated to each treatment group. The block size will not be revealed to any personnel involved in the study except the lead statistician at PRA and the study statistician at Sobi before breaking of the blind.

The first part of the study, i.e. treatment and follow-up of the patients' first flare, is double-blind. Treatment assignment will be blinded for the patients, the investigators and any personnel involved with the study conduct or evaluation at the investigational sites, CRO and sponsor.

Until all patients have completed the Day 15 visit of the first flare (Visit 4), the randomization scheme will only be disclosed to selected CRO personnel to ensure correct packaging of the IMP and correct set-up of the IWRS. The disclosure of the randomization scheme and any individual patient's treatment assignment will be protected by the SOPs of the CRO.

When all patients have completed the Day 15 visit of the first flare, a first database lock will be performed and the main efficacy analysis will be conducted (see Section 6.1). An independent unblinded programmer at the CRO will create unblinded SDTMs to be sent to Sponsor. Blinding will be maintained for patients, investigators, site staff, CRO and applicable vendors throughout the extension phase to prevent any bias in data handling. The Sponsor will take no part in the operational conduct of the study after the first database lock. Details of this process will be outlined in the "Process guideline for data integrity protection after database lock 1".

Unblinding, i.e. breaking the code for an individual patient during the study, is restricted to emergency situations and should only be used under circumstances where knowledge of the treatment is necessary for the proper handling of the patient. The decision to break the code must be made by the investigator. Unblinding should be documented according to instructions in the IWRS manual. The study monitor and sponsor must as soon as possible be informed about the code break.

5 Changes from protocol affecting the statistical analysis

According to the protocol, AEs and SAEs will be classified as occurring during ADA presence when the starting date for the AE is within the time period after the latest negative assessment until next negative assessment with a positive assessment of ADA in between. Instead AEs and SAEs will be presented by ADA positive and ADA negative patients for the whole study period.

The relationship between ADA and efficacy and the relationship between ADA and IL-Ra/anakinra concentrations over time will also be explored which was not described in the protocol.

6 Sequence of planned analysis

6.1 Main efficacy analysis

No interim analysis of efficacy data is planned. However, two database locks are planned. The first database lock will be performed when all patients have conducted the Day 15 visit (Visit 4) of the first flare in the study and the main efficacy analysis will be conducted. The main efficacy analysis includes primarily all efficacy endpoints for the first flare.

All collected data for the first flare, i.e. at a minimum data up to and including the Day 15 visit and at a maximum data up to 12 weeks, will be analyzed. Allocation to analysis populations and all efficacy analyses including data up to Day 15 will be considered as final.

Prior to the first database lock all tasks and criteria defined in the data management plan must be completed and documented. After the database lock the blind will be broken and results will be generated. The SAP will be finalized and signed prior to the first database lock and subsequently the breaking of the blinding. The database lock will be approved by relevant study personnel.

After the data transfer for the main efficacy analysis has occurred, the database will be unlocked to continue with data entry activities for Visit E, Visit 5 and follow-up by phone that had not taken place at the cut-off date for the first database lock and for subsequent flares. However, no changes to the data up to Day 15 should be performed except that ongoing AEs at Day 15 for the first flare might need to be updated.

6.2 Final analysis

The final analysis will include efficacy analyses from subsequent flares and safety and exploratory analyses from all flares. It will be performed when all patients have completed the study. The database will be locked a second time following the same procedure as described in Section 6.1.

If updates to the SAP are needed after the first database lock, all changes will be described and the updated SAP will be finalized, locked and signed prior to final database lock. Any changes to the planned analyses will also be described in the CSR.

6.3 Reporting

The CSR for the study will be written after the final database lock and include data and results from both the first and subsequent flares. The results generated for the main efficacy analysis, i.e. all tables, figures and listings that only includes data up to and including the Day 15 visit, will remain unchanged.

Any post-hoc analyses included in the CSR, which were not identified in the SAP, will be clearly identified as such in the relevant Section of the CSR.

7 Sample size determination

Up to totally 169 patients can be randomized in the study. Initially 159 patients will be randomized, 53 to triamcinolone and 106 to anakinra (53 to 100 mg and 53 to 200 mg). As described in section 4.2, up to a maximum of 10 extra patients can be randomized.

The sample size calculation is based on the change in pain intensity on a VAS 0 to 100 mm from baseline to the average over 24 to 72 hours. A sample size of 106 patients randomized to anakinra and 53 patients randomized to triamcinolone is required to ensure a power of 80 % to reject the null hypothesis of no difference between anakinra and triamcinolone assuming a true difference of 12 mm on VAS mean change and a SD of 25 mm when using a 2-sided test with a significance level of 5 %.

8 Analysis populations

Sections 8.1, 8.2 and 8.3 presents the analysis sets used in the statistical analyses.

8.1 Intention-to-treat (ITT)

The ITT population is the primary analysis population for the primary and secondary efficacy endpoints, comprising all randomized patients grouped according to randomized treatment and stratum. Presentations of efficacy endpoints for subsequent flares will also be based on the ITT population.

8.2 Per-protocol population (PP)

The PP population will comprise all ITT patients who have no major protocol violations potentially affecting the primary efficacy endpoint. The patients will be grouped according to actual treatment received during the first gouty arthritis flare. Patients assigned to an incorrect strata will be grouped according to actual baseline BMI and ULT information. Patient assignment to the PP population will be performed prior to breaking the blind. The PP population will only be used for analyses of efficacy endpoints for the first flare.

The following criteria are classed as major protocol deviations/violations and will exclude patients from the PP population:

- Incorrectly included:
  - Not fulfilling the ACR/EULAR 2015 Gout Classification Criteria.
  - Not fulfilling baseline pain intensity in the index joint of ≥50 on a 0 to 100 VAS or missing baseline pain intensity assessment.
  - The onset of the current ongoing flare is more than 4 days prior to first IMP injection (days from onset to first IMP injection will be calculated in full days allowing patients with time from onset to randomization <5 days to be included in the PP population).
- Missing VAS pain intensity assessments at all of the following timepoints: 24, 48 and 72 hours.
- More than one incorrect intake of rescue medication, i.e. type or outside permitted time window at any of the following timepoints: 24, 48 or 72 hours (see Section 18.1).
- Missing more than one daily IMP injection occasion during the first 3 days of treatment.
- Administration during the first three days of IMP that have had a significant temperature excursion.

8.3 Safety population

The safety population will comprise all patients who received at least one dose of IMP and the patients will be grouped according to actual baseline BMI and ULT information during the first gouty arthritis flare.

Patients randomized to triamcinolone that incorrectly receive at least one dose of anakinra during the first flare will be included in the anakinra group, 100 mg or 200 mg depending on the highest dose received.

Patients randomized to anakinra (100 mg or 200 mg) that incorrectly only receive triamcinolone and no anakinra during the first flare will be included in the triamcinolone group.

The safety population will be used for analyses of safety endpoints and serum concentrations for the whole study period.

The assignment to the safety population will remain the same for all flares even if a patient did change a treatment in a subsequent flare. Changes in treatment for each subsequent flare will be presented in separate tables, i.e. if the actual treatment for the subsequent flare is not the same as the first flare.

9 General considerations for statistical analysis

All statistical tests will be 2-sided and performed using a 5 % significance level, if not stated otherwise. Results will be presented as the point estimate for each treatment group, the estimated difference between groups, the associated 95 %, 2-sided confidence interval and p-value. P-values from statistical analyses will be presented to 3 decimal places with values below 0.001 displayed as <.001 and confidence intervals will be presented to one more decimal place than the raw data.

Continuous data will be summarized using descriptive statistics: n, mean, SD, median, minimum and maximum, unless otherwise stated. Minimum and maximum will be presented to the same number of decimal places as the raw data and mean, SD, standard error of the mean and median will be presented to one more decimal place than the raw data.

Categorical data will be summarized using counts and percentages. Percentages will be suppressed when the count is zero, however the category will still be displayed. The denominator for all percentages will be the number of patients within the treatment group for the population of interest, unless otherwise indicated. Percentages will be presented to one decimal place.

Statistical analyses will be performed using SAS software, Version 9.1 or later (SAS Institute Inc, Cary, North Carolina, United States).

9.1 Handling of missing data and outliers

Patients should continue in the study even if study treatment is withdrawn and every effort should be made to complete the assessments according to the protocol. The evaluation of patient-assessed pain using VAS and Likert scales have repeated assessment time points and repeated measures models will be used to handle missing data, i.e. no imputation is needed. To evaluate the robustness of this model for the primary endpoint, sensitivity analyses including multiple imputation methods will be performed (see Section 14.1.2). For the secondary time-to-event analyses, patients who do not have an event at the time for analysis will be censored at the last

time point with recorded information. For other secondary endpoints the analyses will be based on observed data.

In case of missing or incomplete dates or times, no imputations will be made in the listings of data. Rules on how to handle partially or completely missing dates or times in calculations are described in Section 18.

9.2 Multicenter studies

No effects of centers will be evaluated in this study.

9.3 Multiple comparisons and multiplicity

No adjustment for multiplicity will be performed.

10 Patient disposition

All patient disposition data will be presented by triamcinolone, the combined anakinra group, anakinra 100 mg/day, anakinra 200 mg/day and overall.

The number and percentage of patients who were randomized, initiated treatment, completed and discontinued study together with the reasons for discontinuations will be presented both for each treated flare and for the full study. If a patient prematurely discontinue the study during a flare, all prior flares are completed (i.e. only the last flare may be prematurely discontinued). If the discontinuation is after the Day 15 visit, the flare is considered completed (i.e. all flares may be completed even if the patient has prematurely discontinued the study). Otherwise, the patient has prematurely discontinued the flare with the reason for discontinuation the same as the reason for prematurely discontinuation of the study.

The number of weeks in the study and the number of treated flares experienced by the patients during the study will be presented by descriptive statistics for the ITT and safety population. Further, the number and percentage of patients experiencing 1 treated flare, 2 treated flares, 3 treated flares, etc. will also be presented for the ITT and safety population.

The total number and percentage of patients who completed treatment and discontinued treatment together with reason for discontinuation will be tabulated for each flare and overall for the safety population.

The number and percentage of patients in each strata will be presented for the ITT, PP and safety populations. Stratification data will be received from the IWRS system and actual strata will be derived using CRF data.

The distribution of patients at each site will be tabulated for the ITT population, sorted by descending number of patients.

A table displaying the number and percentage of patients with each type of major protocol deviations that lead to exclusion from the PP population will be presented. Percentages will be based on the ITT population.

The number and percentage of patients in the ITT population, the PP population and the safety population will be presented.

11 Demographics and baseline characteristics

11.1 Demographics

The demographic data, including age, sex, race, and ethnicity will be presented by descriptive statistics for the ITT, PP and safety populations (see Section 18.1).

11.2 Baseline characteristics

Baseline characteristics of gouty arthritis, such as age at diagnosis, disease duration, and number of gouty arthritis flares during the past 12 months will be presented for the ITT, PP, and safety populations (see Sections 18.1 18.4, and 18.5).

The number of patients who completed patient/caregiver IMP injection training and patient diary training will be presented for the ITT population.

The reasons NSAIDs and colchicine are inappropriate will be presented for the ITT, PP and safety populations.

Body height, body weight, BMI, waist circumference, body temperature, systolic and diastolic blood pressure, and pulse (see Section 18.3) will be presented using descriptive statistics for the ITT, PP, and safety populations. For subsequent flares, body weight, BMI, body temperature, systolic and diastolic blood pressure, and pulse will be presented.

Each of the ACR EULAR Gout Classification Criteria at baseline will be presented for the ITT, PP, and safety populations.

Medical examination of current flare (Time of pain onset prior to IMP injection, the number and distribution of joints involved and the index joint) will be presented for each flare for the ITT, PP, and safety populations.

11.3 Medical and surgical history and concurrent illnesses

All medical and surgical history data will be coded using the MedDRA Version 19.1 and the same version will be used throughout the study.

Medical and surgical history at baseline will be presented by SOC and PT for the ITT, PP, and safety populations. Medical and surgical history reported after randomization will be presented separately (see Section 18.6).

Detailed information on the presence of the following selected comorbidities will be presented in a separate table for the ITT, PP, and safety populations: eGFR category, current or previous symptoms of heart failure, diabetes mellitus, hypertension, obesity, cardiovascular disease, hyperlipidemia, cerebrovascular disease and osteoporosis (see Sections **Error! Reference source not found.**, 18.8, 18.9, 18.10 and 18.11).

Data from physical examination will be presented as "normal" or "abnormal" at Day 1 and Day 15 for each flare for the ITT, PP, and safety populations. Any persisting abnormalities should be stated each time the examination is performed in medical history and new abnormalities should be recorded as AEs.

Baseline ECG will be presented for each flare.

12 Prior and concomitant medication

All prior and concomitant medication will be coded using the latest version of World Health Organization drug dictionary (WHODD B2 DEC_2016 DDE [Enhanced]) and the same version will be used throughout the study.

The preferred term grouped by ATC level 4 will be used for presentation and sorted in descending order of frequency in total (all treatment groups together) at the top in the table.

Prior medication, concomitant medication at randomization and onset of new concomitant medication after randomization during the study will be presented separately for the ITT, PP and the safety populations (see Section 18.12).

ULT use, initiated after randomization in addition to the study drug will be presented per flare for patients in stratum ULT use=no for the ITT, PP and the safety populations (see Section 18.14).

13 Study drug compliance

Study drug compliance, calculated as percentage, i.e. actual number of syringes administered divided by number of syringes dispensed times 100 will be presented for each flare and overall by descriptive statistics (see Section 18.15).

14 Efficacy analyses

14.1 Primary efficacy endpoint

The primary efficacy endpoint will be analyzed using the ITT population (primary population for evaluation) and will be based on observed data i.e. no missing data imputation for the primary analysis.

The primary endpoint is the change in patient-assessed pain intensity from baseline to 24 to 72 hours (average of the assessments performed at 24, 48 and 72 hours) for the patients' first flare (see Section 18.16 for time windows for pain intensity assessments).

Patients will score their pain intensity in the joint most affected at baseline (i.e., the index joint) on a 0 to 100 VAS, ranging from no pain (0) to unbearable pain (100) (see Section 18.17).

The comparison of primary interest is between anakinra (100 mg and 200 mg combined) and triamcinolone, and the null and alternative hypotheses with regard to this are defined as:

H_0: $\mu_{\text{anakinra}} = \mu_{\text{triamcinolone}}$

H_A: $\mu_{\text{anakinra}} \neq \mu_{\text{triamcinolone}}$

The primary endpoint will be estimated using a MMRM analysis with the measurements on the individual time points as responses (baseline, 6, 12, 18, 24, 48 and 72 hours), and with treatment (anakinra 100 mg, anakinra 200 mg and triamcinolone), ULT use (yes/no), BMI (<30.0 and ≥30.0 kg/m^2), visit and treatment-visit-interaction as fixed effects and center as random effect. The model is expressed as:

$$y_{ij} = time_j + treat_k + time \times treat_{jk} + ULT_l + BMI_m + center_n + \varepsilon_{ij}$$

where

y_{ij} is the response for patient i on visit j

$time_j$ is the visit effect (baseline, 6, 12, 18, 24, 48 and 72 hours), j=1,…, 7

$treat_k$ is the treatment effect, k=1, 2, 3

ULT_l is the ULT strata effect (yes, no), l=1, 2

BMI_m is the BMI strata effect (<30.0, ≥30.0), m=1, 2

$center_n \sim N(0,\sigma_c^2)$ and $\varepsilon_{ij} \sim N(0,\sigma^2)$

The analysis will be based on a restricted maximum likelihood (REML) approach. The following order of covariance matrices will be used until convergence is met:

i. Unstructured
ii. Heterogeneous Toeplitz
iii. Heterogeneous first order autoregressive
iv. Homogeneous Toeplitz
v. Homogeneous first order autoregressive

vi. Heterogeneous compound symmetry
vii. Homogeneous compound symmetry

Kenward-Roger approximation will be used to estimate denominator degrees of freedom and adjust standard errors.

The estimated mean change from baseline to 24 to 72 hours for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value based on this model will be presented.

The assumptions of normality and homogeneity of variance will be assessed by inspection of normal probability plots and residual plots.

The estimated mean change from baseline to 24 to 72 hours and the associated 95 % confidence interval for anakinra (combined and each dose group) and triamcinolone will also be presented graphically.

14.1.1 Subgroup analyses

In order to examine the consistency of the treatment effect, the MMRM model used for the primary analysis will also be used within the following subgroups, using the ITT population:

- ULT usage (yes/no).
- BMI (<30.0 and ≥30.0 kg/m^2).
- Age (<65 years and ≥65 years).
- Sex (male, female).
- Race (White and Other).
- Renal function (eGFR category ≥90, ≥60 - <90, ≥30 - <60).
- Usage of rescue medication up to 72 hours during for the first flare (yes/no).

For each subgroup, the estimated change in pain intensity in each treatment group (anakinra and triamcinolone), the estimated difference in change in pain intensity between treatments and the corresponding 95 % confidence intervals will be obtained from the model. P-values will not be presented. Further, the change in pain intensity will also be evaluated for the 2 anakinra doses for each subgroup.

The estimated difference in change in pain intensity between anakinra and triamcinolone and the corresponding 95 % confidence intervals will also be presented graphically in a forest plot.

To avoid too small subgroups in the analysis regarding race, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander will be grouped together with Other.

14.1.2 Sensitivity analysis of the primary endpoint

To assess the robustness of the primary analysis results and conclusions on the treatment effect under varying assumptions, sensitivity analyses presented in Sections 14.1.2.1, 14.1.2.2 and 14.1.2.3 will be performed.

14.1.2.1 Sensitivity to subject population

The primary efficacy endpoint will be analyzed using an analysis identical to the primary analysis described in Section 14.1, but for the PP population.

14.1.2.2 Sensitivity to missing data handling

The primary efficacy analysis MMRM approach assumes that the underlying missing data mechanism is MAR. When the assumption of missing data being MAR is valid, the treatment effect estimated from the MMRM analysis is unbiased. However, this assumption may not hold in all circumstances and the possibility that the missingness mechanism is missing not at random (MNAR) should be explored (EMA 2010, National Research Council 2010, Little et al 2002, O'Kelly et al 2014). A sensitivity analysis will be performed to assess the robustness of results from the primary analysis to potential deviations from the MAR assumptions.

In this sensitivity analysis a distinction is made between intermittent missing values and monotone missing values. Intermittent missing values for a variable Y means that some values may be missing, say Y_j is missing, while there are values observed for the same patient at a later time point, i.e. Y_k is observed for some time point $k > j$. Patients with intermittent missing values are said to have a non-monotone missing data pattern. Monotone missing values are missing values where Y_j being missing implies that all Y_k, $k > j$, are missing for that patient. A patient with only monotone missing values (and no intermittent missing values) is said to have a monotone missing data pattern.

In the sensitivity analysis, intermittent missing data are assumed to be MAR, while monotone missing values are assumed to be MNAR in the anakinra treatment groups. The analysis assumes that the missing outcomes on the different treatment arms vary independently, and includes scenarios where monotone missing values in the anakinra treatment arms have worse outcomes than monotone missing values in the triamcinolone treatment arm. Missing data in the triamcinolone group will be assumed to be MAR (intermittent as well as monotone missing values).

The tipping point analysis involves multiple imputation methodology, where each of the imputed datasets is analysed based on the same statistical model used for the primary analysis, i.e. a MMRM analysis of pain intensity VAS at time points baseline, 6, 12, 18, 24, 48 and 72 hours based on the ITT population.

14.1.2.2.1 Exploration of missing data patterns

SAS PROC MI sorts the data into patterns based on whether the variables are missing or observed at each time point, and provides a default output dataset of missing data patterns. Mean pain intensity at each time point will be presented for each missing data pattern by treatment group.

14.1.2.3 Tipping point analysis

The sensitivity analysis will be based on a tipping point approach to assess how severe the departures from the MAR assumption must be in order to overturn a statistical significant result

to a non significant result for the primary analysis. This is done by evaluating several assumptions for the missing data mechanism until the tipping point is reached, i.e. where the study's conclusion changes. The deviation from the MAR assumption is reflected in the shift parameter. If the shift parameter at the tipping point is plausible the conclusions of the primary analysis under the MAR assumption is questionable.

For monotone missing data, the initial shift value will be zero, which represents MAR. The shift will remain at zero for the triamcinolone treatment group. For the anakinra treatment groups, the shift will be increased, and the process repeated until the study's conclusion changes, i.e. treatment effect is no longer significant.

The following steps show the process for the tipping point analysis:

1. Missing VAS values will be imputed based on the Markov-Chain Monte Carlo (MCMC) method (assuming MAR).

   In this step, 500 complete datasets will be generated.

2. For subjects with monotone missing data in the anakinra treatment arms, monotone missing values will be imputed based on a MNAR model by adding a delta value to the MAR imputed value. Delta adjustment will be based on a scaled delta-adjustment approach in which the delta value added at each time point is defined as a percentage of the mean pain VAS at that time point. For the the triamcinolone group monotone missing values, imputed based on the MAR assumption in the first step, will not be adjusted.

   In this step, the 500 datasets with complete data generated in step 1 will be updated.

3. The MMRM model used in the primary analysis will be fitted to the 500 imputed datasets generated in step 2. As in the primary efficacy analysis, the VAS scores from time points baseline, 6, 12, 18, 24, 48 and 72 hours will be included as response. Specification of the model and the contrast of primary interest will match the primary efficacy analysis. From each analysis of an imputation dataset, adjusted estimates of means, mean treatment differences and p-values will be obtained.

4. The results from the MMRM analyses of the 500 imputed datasets will be combined to obtain an overall estimate of the contrast of primary interest, i.e. the mean difference between the anakinra groups and the triamcinolone group in change from baseline to 24-72 hours in pain VAS score, as well as the p-value obtained after accounting for the uncertainty introduced by the missing data.

5. Steps 2 to 4 will be repeated by gradually increasing the values for the scaled delta adjustment until the scaled delta value is found that gives a p-value exceeding 0.05 (2-sided) or greater.

In the first step of the delta-adjustment, missing values are imputed based on a MCMC method assuming MAR. The MCMC sampling will be performed as implemented in PROC MI in SAS. The MCMC method will include a single chain with a burn-in of 1000, a thinning of 100 and non-informative priors for all parameters.

Treatment group and strata (ULT usage and BMI) are included in the variable list in the MCMC step. In this implementation, numeric dummy variables are created and included in the variable list together with the VAS pain intensity variables for time points 0, 6, 12, 18, 24, 48 and 72 hours. Because these dummy variables are observed for all patients, they are not imputed in the MCMC step, but contribute to the imputation model to allow for the mean pain intensity at different time points to depend on treatment and strata, while assuming the covariance structure across treatment and strata. This corresponds to the assumptions made in the analysis model used for the primary analysis.

In the second step, monotone missing values in the anakinra treatment arms are delta-adjusted where a delta value is added to the MAR imputed value. This approach corresponds to the third approach for pattern imputation with delta adjustment described in Ratitch et al. (2013).

The delta is defined in terms of percentage of the mean value at a given time point. Because the mean VAS pain intensity changes with time, the delta added to the MAR-imputed missing values will be different for different time points. Imputed values will be truncated at 100, since this is the maximum value possible for a VAS pain intensity measure. No rounding of values is performed in the imputation step.

In the third step, the complete datasets are analysed based on the same analysis model used for the primary analysis. In the fourth step, results across the different imputed datasets are combined based on the methodology of Rubin (1976) as implemented in PROC MIANALYZE.

Five hundred (500) imputations will be performed by using a random seed generated from the system clock in SAS. In the search for the tipping point, increasing values for the scaled delta will be investigated in steps of 5% of the time-specific mean pain intensity VAS.

In the fifth step, when potentially a tipping point has been found, LS Mean estimates and treatment differences with 95 % confidence intervals and p-value will be presented in a table for the choice of delta-adjustment that tip over the conlusions.

Since the tipping point analysis aims to find out how much of a change is required to “tip” the result from statistically significant to not statistically significant, it will not be performed for any comparison that is not statistically significant at the onset.

14.2 Secondary endpoints supporting the primary objective

The secondary endpoints supporting the primary objective will be analyzed using both the ITT population and the PP population.

14.2.1 Patient-assessed pain intensity in the index joint on a Visual analogue scale (VAS)

To further evaluate the effect of anakinra, the change from baseline in pain intensity on VAS at 6, 12, 18, 24, 36, 48, 72 hours and at Days 5, 6, 7 and 8 for the first flare will be evaluated.

Pain intensity on VAS will be evaluated using a mixed model repeated measures analysis similar to the primary analysis with the measurements on the individual time points as responses (baseline, 6, 12, 18, 24, 36, 48 and 72 hours and Days 5, 6, 7 and 8), and with treatment (anakinra 100 mg, anakinra 200 mg and triamcinolone), ULT use (yes/no), BMI (<30.0 and ≥30.0 kg/m^2), visit and treatment-visit-interaction as fixed effects and center as random effect. For each time point, the estimated mean change from baseline for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value based on this model will be presented.

The same order of covariance matrices as for the primary endpoint will be used, until convergence is met.

Graphs will be used to visualize the mean pain intensity on VAS estimated from the repeated measures model at each time point for each dose group of anakinra and triamcinolone and also by subgroups of BMI and ULT usage.

The observed patient-assessed pain intensity and change from baseline (VAS) in index joint at each time point will also be presented by descriptive statistics.

In addition, the percent change in observed pain intensity at 24, 48 and 72 hours will be visualized by graphically presenting the distribution of percent improvement. For each category of percent improvement (x-axis), the percentage of patients who improved (y-axis) will be plotted. Increments of 10 % will be used for the improvement (≥10 %, ≥20 %, ≥30 %, etc).

14.2.2 Patient-assessed pain intensity in the index joint on a 5-point Likert scale

To further evaluate the effect of anakinra, the change from baseline in pain intensity on the Likert scale at 6, 12, 18, 24, 36, 48, 72 hours and at Days 5, 6, 7 and 8 for the first flare will be evaluated (see Section 18.18).

Pain intensity on the Likert scale will be evaluated using a mixed model repeated measures analysis similar to the primary analysis with the measurements on the individual time points as responses (baseline, 6, 12, 18, 24, 36, 48 and 72 hours and Days 5, 6, 7 and 8), and with treatment (anakinra 100 mg, anakinra 200 mg and triamcinolone), ULT use (yes/no), BMI (<30.0 and ≥30.0 kg/m^2), visit and treatment-visit-interaction as fixed effects. For each time point, the estimated mean change from baseline for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value based on this model will be presented.

The same order of covariance matrices as for the primary endpoint will be used, until convergence is met.

Graphs will be used to visualize the mean pain intensity on the Likert scale estimated from the repeated measures model at each time point for each dose group of anakinra and triamcinolone.

The observed patient-assessed pain intensity and change from baseline (Likert) as well as observed patient-assessed pain intensity categories (Likert) in index joint at each time point will also be presented by descriptive statistics.

14.3 Secondary efficacy endpoints

The secondary efficacy endpoints for the 1st flare will be analyzed using both the ITT population and the PP population.

14.3.1.1 Pain intensity for anakinra dose groups, 100 and 200 mg

The primary endpoint, change in patient-assessed pain intensity as measured by VAS from baseline to 24 to 72 hour for the first flare, will be evaluated for the 2 different anakinra doses (100 mg and 200 mg) in comparison to triamcinolone.

The effect of the 2 different anakinra doses will be evaluated using the same mixed model repeated measures model analysis as for the primary endpoint. The estimated mean change from baseline to 24 to 72 hours for each of the 2 anakinra doses (100 mg and 200 mg) and triamcinolone, the estimated differences anakinra 100 mg – triamcinolone and anakinra 200 mg – triamcinolone and the associated 95 % confidence intervals and p-values from the model will be presented.

The estimated mean change from baseline for the secondary endpoints (VAS and Likert pain intensity at each timepoint) supporting the primary endpoint for each of the 2 anakinra doses (100 mg and 200 mg) and triamcinolone, the estimated differences anakinra 100 mg – triamcinolone and anakinra 200 mg – triamcinolone and the associated 95 % confidence intervals and p-values from the model will also be presented.

14.3.1.2 Onset of effect

Onset of effect is defined as at least 20 % reduction from baseline pain intensity on VAS, and the time to onset of effect for the first flare will be analyzed using a stratified log-rank test, with ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2) as stratification factors (see Section 18.19). Based on this model, the estimated hazard ratio (anakinra versus triamcinolone), 95 % confidence interval and p-value will be presented. Further, the hazard ratios for the comparisons of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values.

The hazard ratio and its confidence interval will be estimated based on Peto's method:

$HR = \exp(U/V)$

95 % CI for HR = $(\exp\{U/V - 1.96/\sqrt{V}\}, \exp\{U/V + 1.96/\sqrt{V}\})$

Where $U = \sum (d_{1i}-e_{1i})$ is the log-rank test statistic (with d_{1i} and e_{1i} the observed and expected events in group 1) and $\sqrt{V}$ the SD of the log-rank test statistic as produced in the PROC LIFETEST output. This approach will be adapted for a stratified approach by using the U and V statistics obtained directly from the LIFETEST having added a STRATA term for each of the

stratification variables. Patients with only baseline values and patients with no values will be censored at baseline, i.e. will not be included in the analysis.

Time to onset of effect will also be presented graphically for the two dose groups of anakinra and triamcinolone.

In addition, the number and proportion of patients with onset of effect in the different treatment groups up to 6, 12, 24, 48 and 72 hours will be presented. The percentages will be based on the number of patients in the ITT population.

The number and percentage of patients with and without onset of effect within 24 hours versus the number and percentage of patients with and without response within 72 hours will also be tabulated.

14.3.1.3 Response

Response is defined as at least 50 % reduction from baseline pain intensity on VAS, and the time to response for the first flare will be analyzed using a stratified log-rank test, with ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2) as stratification factors (see Section 18.20). Based on this model, the estimated hazard ratio (anakinra versus triamcinolone), 95 % confidence interval and p-value will be presented. Further, the hazard ratios for the comparisons of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values.

The hazard ratio and its confidence interval will be estimated based on Peto's method (see Section 14.3.1.2). Patients with only baseline values and patients with no values will be censored at baseline, i.e. will not be included in the analysis.

Time to response will also be presented graphically for the two dose groups of anakinra and triamcinolone.

In addition, the number and proportion of responders in the different treatment groups at 24, 48 and 72 hours, Day 8 and Day 15 will be presented. The percentages will be based on the number of patients in the ITT population.

The number and percentage of patients with and without response within 72 hours versus the number and percentage of patients with and without onset within 24 hours will also be tabulated.

14.3.1.4 Resolution of pain

Resolution of pain is defined as <10 mm in patient-assessed pain intensity on VAS, and the time to resolution for the first flare will be analyzed using a stratified log-rank test, with ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2) as stratification factors (see Section 18.21). Based on this model, the estimated hazard ratio (anakinra versus triamcinolone), 95 % confidence interval and p-value will be presented. Further, the hazard ratios for the comparisons of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values.

The hazard ratio and its confidence interval will be estimated based on Peto's method (see Section 14.3.1.2). Patients with only baseline values and patients with no values will be censored at baseline, i.e. will not be included in the analysis.

Time to resolution of pain will also be presented graphically for the two dose groups of anakinra and triamcinolone.

In addition, the number and proportion of patients with resolution of pain in the different treatment groups at 72 hours, Day 8 and Day 15 will be presented. The percentages will be based on the number of patients in the ITT population.

14.3.1.5 Rescue medication

Time to first intake of rescue medication during the first flare will be analyzed using a stratified log-rank test, with ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2) as stratification factors (see Section 18.22). Based on this model, the estimated hazard ratio (anakinra vs. triamcinolone), 95 % confidence interval and p-value will be presented. Further, the hazard ratios for the comparisons of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values.

The hazard ratio and its confidence interval will be estimated based on Peto's method (see Section 14.3.1.2). Patients with only baseline values and patients with no values will be censored at baseline, i.e. will not be included in the analysis.

Time to first intake of rescue medication will also be presented graphically for the two dose groups of anakinra and triamcinolone.

In addition, the proportion and number of patients who took rescue medication from the first IMP administration up to 12 hours, >12 to 24 hours, >24 to 48 hours, >48 to 72 hours, >72 hours to Day 8, and >Day 8 to Day 15 in the different treatment groups will be presented. The proportion and number of patients with 0, 1, 2, 3 or more than 3 occasions with rescue medications as well as descriptive statistics for number of occasions will be presented in the same time intervals. Multiple occurences of rescue medication within 15 minutes will only count as one occurence.

The type of rescue medication taken will also be summarized in the same time intervals per flare.

14.3.1.6 Patient´s assessment of global response to treatment

Patient's assessment of global response to treatment at 72 hours, Day 8 and Day 15 for the first flare will be evaluated using an analysis of variance including factors for treatment (anakinra 100 mg, anakinra 200 mg and triamcinolone), ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2). The estimated mean global response to treatment for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value from the model will be presented. Further, the estimated mean global response for each of the 2 anakinra doses (100 mg and 200 mg), the difference of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values.

The observed patient-assessed global response categories at each time point will also be presented by descriptive statistics.

The cumulative proportion of patients that reach good patient's assessed global response to treatment (where good response is defined as good, very good or excellent) over time will be illustrated graphically.

14.3.1.7 Physician's assessment of global response to treatment

Physician's assessment of global response to treatment at 72 hours, Day 8 and Day 15 for the first flare will be evaluated using an analysis of variance model including factors for treatment (anakinra 100 mg, anakinra 200 mg and triamcinolone), ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2). The estimated mean global response to treatment for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value from the model will be presented. Further, the difference of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values.

The observed physician-assessed global response categories at each time point will also be presented by descriptive statistics.

The cumulative proportion of patients that reach good physician's assessed global response to treatment (where good response is defined as good, very good or excellent) over time will be illustrated graphically.

14.3.1.8 Physician's assessment of clinical signs

The investigator will assess the tenderness, swelling and presence of erythema in the index joint.

Physician's assessment of tenderness at 72 hours, Day 8 and Day 15 for the first flare will be evaluated using an analysis of covariance model including factors for treatment (anakinra 100 mg, anakinra 200 mg and triamcinolone), ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2), and baseline tenderness as covariate. The estimated mean tenderness for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value from the model will be presented. Further, the estimated mean tenderness for each of the 2 anakinra doses (100 mg and 200 mg), the difference of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values. The observed physician-assessed clinical signs categories at each time point will also be presented by descriptive statistics.

Physician's assessment of swelling at 72 hours, Day 8 and Day 15 for the first flare will be evaluated using an analysis of covariance model including factors for treatment (anakinra 100 mg, anakinra 200 mg and triamcinolone), ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2), and baseline swelling as covariate. The estimated mean swelling for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value from the model will be presented. Further, the estimated mean swelling for each of the 2 anakinra doses (100 mg and 200 mg), the difference of the 2 different

anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values. The observed physician-assessed swelling categories at each time point will also be presented by descriptive statistics.

The presence of erythema at 72 hours, Day 8 and Day 15 for the first flare will be analyzed using a logistic regression model with treatment, ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2) as explanatory variables. The proportion of patients with presence of erythema in each treatment group, the estimated odds ratio of anakinra to triamcinolone, the corresponding 95 % confidence interval and the p-value from the model will be presented. Further, the odds ratios for the comparisons of the 2 different anakinra dose groups versus triamcinolone will also be estimated, with corresponding 95 % confidence intervals and p-values. The observed physician-assessed presence of erythema at each time point will also be presented by descriptive statistics.

14.3.1.9 Inflammatory biomarkers

The change from baseline in CRP (using the central laboratory value) at 72 hours, Day 8 and Day 15 will be evaluated using an analysis of covariance model including factors for treatment, ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2), and baseline CRP as covariate. The estimated mean change in CRP for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value from the model will be presented. Further, the estimated mean change in CRP for each of the 2 anakinra doses (100 mg and 200 mg), the estimated differences anakinra 100 mg – triamcinolone and anakinra 200 mg – triamcinolone, and the associated 95 % confidence intervals and p-values from the model will be presented. The observed CRP values at each time point will also be presented by descriptive statistics.

The change from baseline in SAA at 72 hours, Day 8 and Day 15 will be evaluated using an analysis of covariance model including factors for treatment, ULT use (yes, no) and BMI (<30.0, ≥30.0 kg/m^2), and baseline SAA as covariate. The estimated mean change in SAA for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value from the model will be presented. Further, the estimated mean change in SAA for each of the 2 anakinra doses (100 mg and 200 mg), the estimated differences anakinra 100 mg – triamcinolone and anakinra 200 mg – triamcinolone, and the associated 95 % confidence intervals and p-values from the model will be presented.

The observed SAA values at each time point will also be presented by descriptive statistics.

The evaluation of CRP and SAA at Week 12 will be presented by descriptive statistics.

14.3.1.10 Subsequent flares

The evaluation of efficacy at subsequent flares will be presented by the combined anakinra group, the 2 different anakinra doses (100 mg and 200 mg) and triamcinolone, and by flare, i.e. patient's 2nd flare, 3rd flare, etc. for the ITT population.

Change in patient-assessed pain intensity in the index joint as measured by VAS score from baseline to 24 to 72 hours (average of the assessments performed at 24, 48 and 72 hours) will be estimated using the same mixed model repeated measures analysis as for the primary endpoint.

The estimated mean change from baseline to 24 to 72 hours for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval will be presented for each flare. Graphs will be used to visualize pain intensity (VAS) at each time point for anakinra (combined and each dose group) and triamcinolone estimated from the repeated measures model for each flare. In addition, graphs for each treatment group will be used to visualize pain intensity (VAS) over time for all flares in the same graph.

Patient-assessed pain intensity on VAS will be summarized by descriptive statistics for each time point, i.e. baseline, 6, 12, 18, 24, 36, 48, and 72 hours and Days 5, 6, 7 and 8, as well as the change from baseline at these time points. The presentations will be repeated for patients 2nd flare, 3rd flare, etc.

Patient-assessed pain intensity on Likert will be summarized by descriptive statistics for each time point, i.e. baseline, 6, 12, 18, 24, 36, 48, 72 hours and Days 5, 6, 7 and 8 as well as the change from baseline at these time points. The presentations will be repeated for patients 2nd flare, 3rd flare etc.

Further, the mean pain intensity on both VAS and Likert over time will be visualized graphically. For each treatment group, the mean pain intensity over time will be displayed for each flare.

The number and proportion of responders in the treatment groups at 24, 48 and 72 hours, Day 8 and Day 15 will be summarized by descriptive statistics. The percentage will be based on the number of patients in the ITT population that experienced the specific number of flares.

The number and proportion of patients with onset of effect in the different treatment groups at 6, 12, 24, 48 and 72 hours will be summarized by descriptive statistics. The percentage will be based on the number of patients in the ITT population that experienced the specific number of flares.

The number and proportion of patients with resolution of pain in the treatment groups at 72 hours, Day 8 and Day 15 will be summarized by descriptive statistics. The percentage will be based on the number of patients in the ITT population that experienced the specific number of flares.

The number and proportion of patients that took rescue medication from the first IMP administration at Visit 1 up to and including the Day 15 visit in the treatment groups will be summarized by descriptive statistics. The proportion and number of occasions of rescue medication events and the type of rescue medication will also be presented.

Patient’s assessment of global response to treatment at 72 hours, Day 8 and Day 15 will be summarized by descriptive statistics.

Physician’s assessment of global response to treatment at 72 hours, Day 8 and Day 15 will be summarized by descriptive statistics.

Physician’s assessment of clinical signs (tenderness, swelling and presence erythema) at 72 hours, Day 8 and Day 15 will be summarized by descriptive statistics.

The change from baseline in CRP and SAA at 72 hours, Day 8 and Day 15 will be summarized by descriptive statistics.

15 Serum concentrations of Interleukin 1 receptor antagonist (IL-1Ra)/anakinra

The serum concentration of endogenous IL-1Ra/anakinra at baseline, Day 8, Day 15, Day 28 and Week 12 for the first flare and subsequent flare will be presented by descriptive statistics.

16 Safety analyses

The safety population is based on actual treatment during first flare and there is a risk that patients´ treatment at new flares differs from the first flare. Therefore the number of patients and the actual treatment for each of the subsequent flare will be tabulated versus the number of patients in the treatment groups assigned for the safety population.

16.1 Drug exposure

Treatment exposure, calculated as total amount anakinra/triamcinolone given per flare in mg will be presented for each flare and for the whole study by descriptive statistics (see Section 18.23).

16.2 Adverse events

All analyses of AEs will be based on the safety population. No hypothesis testing will be performed. AEs will be coded using MedDRA Version 19.1. Percentages will be based on the number of patients in the safety population for the specific treatment group.

Since a subsequent flare can occur before the AE reporting period for the current flare has ended, the actual length of the AE reporting period can differ between patients. The actual length of the AE and the SAE reporting period will be presented by descriptive statistics for each flare and for the whole study.

A TEAE is defined as an AE with a start date/time after the first administration of IMP and not after the last day of the AE reporting period for each flare. AEs with unknown start date/time will be assumed to be treatment emergent unless the end date/time is known to be before the first administration of IMP. Pre-treatment AEs will be identified as AEs with a start date/time before the first administration of IMP (see Section 18.24).

The number and percentage of patients with at least one TEAE, at least one severe TEAE, at least one serious TEAE, including death, at least one non-serious TEAE, any related TEAE, any fatal TEAE, any TEAE leading to drug withdrawn and any TEAE leading to study withdrawal will be summarized.

The number and percentage of patients with at least one TEAE will be summarized in frequency tables by SOC and PT. This table will be repeated by BMI (<30.0, ≥30.0 kg/m^2), sex, age (<65, ≥65 years), race, and by ethnic origin for each flare and overall.

The number and percentage of patients with at least one TEAE will be summarized in frequency tables separately by day of onset: Day 1 to 5, Day 6 to 15 and after Day 15 by SOC and PT for each flare and overall.

The number and percentage of patients with at least one TEAE will be summarized in frequency tables by PT in descending order of frequency in the combined anakinra group for each flare and overall.

The number and percentage of patients with at least one TEAE will be summarized in frequency tables by PT, classified by maximum severity, in descending order of frequency in the combined anakinra group for each flare and overall.

The number and percentage of patients with at least one TEAE will be summarized in frequency tables by PT, classified by relationship in descending order of frequency in the combined anakinra group for each flare and overall.

The number and percentage of patients with non-serious TEAE will be presented by SOC and PT in descending order of frequency in the combined anakinra group for each flare and overall.

16.2.1 Serious adverse events

The number and percentage of patients with serious TEAE will be presented by SOC and PT in descending order of frequency in the anakinra group for each flare and overall. The narratives will be presented in listings.

The number and percentage of patients with at least one pre-treatment SAE will be summarized in frequency tables by SOC and PT.

The number and percentage of patients with at least one SAE will be summarized in frequency tables separately by day of onset: Day 1 to Day 28 and after Day 28 for each flare and overall.

16.2.2 Adverse events leading to withdrawal

The number and percentage of patients with TEAEs leading to study withdrawal will be presented by SOC and PT, in descending order of frequency in the anakinra group for each flare and overall.

The number and percentage of patients with TEAEs leading to drug withdrawal will be presented by SOC and PT, in descending order of frequency in the anakinra group for each flare and overall.

16.2.3 Deaths and other serious adverse events

A summary of deaths and other serious adverse events will be presented in listings.

The narratives of deaths and other serious adverse events will also be presented in listings.

16.3 Laboratory data

The presentations will be based on the results from the central laboratory safety data.

The laboratory safety data will be presented as actual values at baseline, 72 hours, Day 8 and Day 15 and changes from baseline at 72 hours, Day 8 and Day 15 for each flare by descriptive statistics.

The number and percentage of patients with low, normal, or high laboratory values at baseline versus 72 hours, Day 8 and Day 15 for the first flare will be presented using shift tables. For subsequent flares the number and percentage of patients with low, normal, or high laboratory values at baseline for the flare of interest versus Day 15 will be presented using shift tables.

The values at 72 hours, Day 8 and Day 15 versus baseline for the flare of interest for each laboratory parameter for each flare will be presented in a scatterplot. Side by side plots for anakinra (different symbols for 100 mg and 200 mg) and triamcinolone will be performed.

Treatment emergent abnormal laboratory values will be listed. Values that are normal or high at baseline for each flare and shift to low at the corresponding visit, or that are normal or low at baseline and shift to high at the corresponding visit.

16.4 Vital signs

Vital signs (blood pressure, pulse rate and body temperature) will be presented as actual values at baseline, 72 hours, Day 8 and Day 15 for each flare by descriptive statistics.

Treatment emergent abnormal vital signs will be listed. Values that are normal or high at baseline for each flare and shift to low at the corresponding visit, or that are normal or low at baseline and shift to high at the corresponding visit will be listed. See Section 18.25 for definition of low, normal and high.

16.5 Electrocardiogram

ECG results at baseline will be summarized for each flare by descriptive statistics.

16.6 Anti-drug antibodies

See Section 18.25 for definitions regarding immunogenicity.

16.6.1 Presence of anti-drug antibodies

The number and proportion of patients with presence of ADA (treatment induced, treatment boosted or treatment unaffected), NAb and cross-reactivity at baseline, Day 8, Day 15, Day 28 and Week 12 will be summarized for each flare.

For patients with presence of ADA, the ADA titers at baseline, Day 8, Day 15, Day 28 and Week 12 will be summarized by descriptive statistics for each flare. Depending on how the results are reported, ADA titers might be categorized as eg. low (<50), medium (>50 and <500) or high (>500).

The cumulative number and proportion of ADA positive and ADA negative subjects will also be summarized by flare over the study.

The number and proportion of transient ADA positive patients, persistent ADA positive patients and off-treatment persistent ADA positive patients will be summarized for the whole study,

The number of patients with occurrence of transient and persistent ADA during the study will be presented graphically.

16.6.2 Anti-drug antibodies versus safety

AEs and SAEs will be presented by ADA positive and ADA negative patients for the whole study period by SOC and PT. ADA positive patients will then be further categorized in:

1. Transient ADA positive patients.
2. Persistent ADA positive patients.

SAEs will be listed together with the ADA titer for the positive ADA sample taken closest before and after the start of the event.

16.6.3 Anti-drug antibodies versus efficacy

The mean pain intensity on VAS at each timepoint up to Day 8 will be presented by ADA presence and by NAb presence by descriptive statistics and in graphs as applicable. The presentations will be repeated for each subsequent flare. ADA presence will be defined as being ADA positive at least once at Day 1 or Day 8; no ADA presence will be defined as being negative at both Day 1 and Day 8. Undetermined ADA presence is defined as both assessments are missing or if one assessment is negative and the other is missing.

For subsequent flares, the ADA titer at Day 1 will be plotted against VAS at 72 hours.

For baseline, Day 8 and Day 15, CRP and SAA will be presented by ADA presence and by NAb presence. The presentations will be repeated for each subsequent flare. ADA presence will be defined as being ADA positive at least once at Day 1, Day 8 or Day 15; no ADA presence will be defined as being negative at both Day 1, Day 8 and Day 15. Undetermined ADA presence is defined as all assessments are missing or if 1 or 2 assessments are negative and the other assessments are missing.

For subsequent flares the ADA titer will be plotted against CRP and SAA at Day 1, Day 8 and Day 15.

16.6.4 Anti-drug antibodies versus serum concentration of IL-1Ra/anakinra

If ADA are present, the relationship between ADA and ADA titer and IL-Ra/anakinra concentrations over time will be explored, as applicable. Similarly the relation between presence of NAb and IL-Ra/anakinra concentrations over time will be explored.

17 Exploratory endpoints

17.1 Short Form (36) Health Survey (SF-36)

Based on the SF-36 questionnaire, the patient's score for each of the individual 8 domains as well as the PCS and the MCS will be calculated (see Section 18.27).

The change from baseline in SF-36 physical functioning domain score at Day 8 for the first flare will be evaluated using an analysis of covariance model including factors for treatment, ULT use (yes, no) and BMI (<30.0, ≥ 30.0 kg/m^2), and baseline SF-36 physical functioning score as covariate. The estimated mean change in SF-36 physical functioning domain score for each treatment group (anakinra and triamcinolone), the estimated difference between the groups and the associated 95 % confidence interval and p-value from the model will be presented.

The patient's score and the change from baseline for the 8 domains as well as the PCS and the MCS at Day 8 and Day 15 will be presented by descriptive statistics.

For subsequent flares, the 8 domains from SF-36 as well as the PCS and the MCS at Day 8 and Day 15 will be presented by descriptive statistics.

17.2 EuroQol 5 dimensions 5 levels (EQ-5D-5L)

The actual values and the changes from baseline in the 5 domains in EQ-5D, the EQ-5D-VAS and the EQ-5D Index at baseline, Day 8 and Day 15 for first and subsequent flares will be presented by descriptive statistics (see Section 18.28).

The EQ-5D Index value will also be presented by resolution of pain until Day 8 for the first flare.

In addition the distribution of categories in each EQ-5D domain will be presented at baseline and Day 8.

17.3 Work productivity and activity impairment:specific health problems (WPAI:SHP)

Based on the WPAI:SHP the following scores will be calculated for Day 8 and Day 15; percent and actual work time missed due to gout, percent impairment while working due to gout, percent overall work impairment due to gout, percent activity impairment due to gout and in addition total work time will be calculated (see Section 18.29).

The WPAI:SHP scores at Day 8 and Day 15 for the first flare and subsequent flares will be presented by descriptive statistics.

17.4 Health care resource utilization

The number of days with hospitalization and the number of unscheduled outpatient visits (Specialist, Nurse practitioner and GP/Family Physician) from Day 1 to Day 15 for the first flare and subsequent flares will be presented by descriptive statistics (see Section 18.30).

17.5 Exploratory inflammatory biomarkers

The actual values at all timepoints and change from baseline in exploratory inflammatory biomarkers in serum (IL-6, IL-8, calprotectin, MPO and neopterin) at 72 hours, Day 15 and Week 12 (in case of no subsequent flare prior to the Week 12 time point) for the first flare and subsequent flares will be presented by descriptive statistics.

18 Derived and computed variables

18.1 Unpermitted rescue medication

Unpermitted rescue medication that leads to exclusion from PP population is defined as more than one incorrect intake, i.e. either type or outside time window (multiple incorrect intakes within 15 minutes is counted as one):

- Unpermitted type
    - Types other than paracetamol/acetaminophen and/or codein or tramadol or rescue medications classified as non-medicinal products (e.g. ice/cold pack, vinegar) and taken between treatment start and 72 hours.
    - Glucocorticoids (from general concomitant medication and rescue medication) taken between treatment start and 72 hours).
    - All unresolved OTHER <72 hours.

- Outside time window
    - All rescue (from rescue medication) taken prior to the 24 hours VAS pain assessment. If 24 hours VAS pain is missing, taken within 24 hours since treatment start.
    - All rescue (from rescue medication) taken within 6 hours prior to the 48 hours VAS pain assessment. If 48 hours VAS pain is missing, taken within 42 to 48 hours since treatment start.
    - All rescue (from rescue medication) taken within 6 hours prior to the 72 hours VAS pain assessment. If 72 hours VAS pain is missing, taken within 66 to 72 hours since treatment start.
    - All rescue (from rescue medication) with a missing timestamp (date and time) that can be related to a flare.

18.2 Age

Age will be calculated in full years without decimals, derived through the formula: int((input(compress(RFSTDTC, '-'), best.) - input(compress(BRTHDTC, '-'), best.))/10000). The method automatically corrects for leap days but is only correct for values without decimals.

18.3 Body mass index (BMI)

BMI will be calculated as:

*BMI = weight (kg)/(height (m)*height (m))*

18.4 Age at diagnosis of gouty arthritis

Age at diagnosis of gouty arthritis will be calculated in full years without decimals, derived through the formula: int((input(compress(MHSTDTC (where MHTERM = GOUTY ARTHRITIS) , '-'), best.) - input(compress(BRTHDTC, '-'), best.))/10000).

Since the date of diagnosis of gouty arthritis is only collected as month and year, the day will be imputed as the 15th, i.e. YYYY-MM-15, in the calculation. If the month is missing, it will be imputed as June, i.e. YYYY-06-15.

18.5 Disease duration

The disease duration (calculated in full years) is defined as:

- If the date of diagnosis is given only as a year: year of randomization - year of diagnosis.
- If the date of diagnosis is given as a year and a month: INT((YYYYMM$_{randomization}$ - YYYYMM$_{diagnosis}$)/100).

18.6 Start of medical and surgical history

If the start date is missing, the medical and surgery history will be reported as started after randomization (first IMP injection).

18.7 eGFR category

The definition of eGFR categories is presented in Table 1.

Table 1 eGFR categories

eGFR categories (mL/min/1.73 m^2)
≥90
60 to 89
30 to 59
15 to 29
<15

Abbreviations: GFR, Glomerular filtration rate.

The eGFR will be calculated according to the MDRD (Modification of Diet in Renal Disease) formula:

eGFR (mL/min/1.73 m^2) = 175 × $(S_{cr})^{-1.154}$ × $(Age)^{-0.203}$ × (0.742 if female) × (1.212 if African American).
Where S_{cr} is serum creatinine in µmol/L.

18.8 Obesity

Obesity is defined as a BMI ≥30.

18.9 Cardiovascular disease

Cardiovascular disease is defined as a history of angina pectoris or myocardial infarction.

18.10 Hyperlipidemia

Hyperlipidemia is defined as Cholesterol >240 mg/dL or taking stable cholesterol-lowering treatments defined as ATC codes: C10, increased low-density lipoprotein levels (>160 mg/dL) , hypertriglyceridaemia (triglycerides >150 mg/dL), or decreased high-density lipoprotein levels (<40 mg/dL in male or <50 mg/dL in female).

18.11 Cerebrovascular disease

Cerebrovascular disease is defined as previous ischemic or hemorrhagic stroke, previous TIA, carotid or vertebral stenosis/occlusion.

18.12 Start of prior and concomitant medication

If the start or end date of medication is (partly) missing so that categorization in time periods of medication cannot be done, the question regarding if medication was taken prior to the study will be used for the categorization.

18.13 Glucocorticoid use

Glucocorticoid use is defined as ATC codes: H02A, H02B or M01BA.

18.14 ULT use

ULT use during the study is defined as ATC codes: M04AA, M04AB or M04AX.

18.15 Study drug compliance

The first IMP injections will be given at site (one i.m. and 2 s.c. injections) at Visit 1 and drugs for 2 days (24 and 48 hours) will be dispensed. At Visit 2 (72 hours) the 2 s.c. injections will be given at site and drug will be dispensed for one additional day (Day 5). The number of syringes given at site, the number of days for which IMP was dispensed (2 syringes per day) and number of syringes returned will be recorded in the CRF. The number of syringes used at home will be recorded in the eDiary at 24 and 48 hours and at Day 5.

Drug compliance will be calculated as:

> ([Number of syringes dispensed – number of syringes returned]/ number of syringes dispensed)*100

Number of syringes dispensed will be calculated as:

> Total number of syringes supposed to be given at site at Visit 1 (3 syringes) and Visit 2 (2 syringes) + total number of syringes dispensed at Visit 1 and Visit 2

Number of syringes returned will be calculated as:

> Number of unused syringes at site at Visit 1 and Visit 2 + total number of unused syringes returned

18.16 Time windows for pain intensity assessments (Visual analogue scale – VAS- and Likert)

The time windows presented in Table 2 are used for collecting pain assessments in the eDiary at the nominal times for baseline, 6, 12, 18, 24, 36, 48, 72 hours and Days 5, 6, 7,8, 9, 10, 11, 12, 13, 14 and 15. Time for first IMP injection is time=0. For assessments where the nominal time would fall during the patients sleep, pain intensity was collected prior to sleep and at awakening. These assessments will be assigned to the nominal time if they fall within its below time-window.

Table 2 Time windows for pain intensity assessments (VAS and Likert)

	Time window	
Nominal time	**Start time (h)**	**End time (h)**
Baseline		≤0
6 h	>0	≤9
12 h	>9	≤15
18 h	>15	≤21
24 h	>21	≤30
36 h	>30	≤42
48 h	>42	≤60
72 h	>60	≤84
Day 5	>84	≤108
Each following day until Day 15 but day 15 does not have an end.	+24	+24

Abbreviations: h, hours; VAS, Visual analogue scale.

If IMP is not taken the following rules will be used:

- For the first flare, for patients not receiving treatment the baseline VAS pain measurement from their date of randomization will be used as the baseline assessment.
- For subsequent flares, if a patient doesn't receive treatment, the flare will not be accounted for.

If multiple assessments fall within the same windows the following rules will be used:

- If multiple assessments fall into the same time window, the one closest to the nominal time will be used.
- If multiple assessments in the same time window are equidistant to the nominal time, the latest will be used.
- If multiple assessments have the exact timestamp, an average will be used.

18.17 Pain intensity on Visual analogue scale (VAS)

The patients will score their current pain intensity in the joint most affected (i.e., the index joint) on a continuous 0 to 100 VAS, ranging from no pain (0) to unbearable pain (100). The original 100 mm scale has been transferred into ePRO format and proportionally scaled down to 5 cm in length. The conceptual equivalency of the pen-and-paper and the ePRO format has been tested and is described in a Cognitive Debriefing and Usability Testing of Pain Scales for Gouty Arthritis Report.

For each flare, the pain intensity will be assessed in the index joint and recorded in the patient diary prior to the first IMP administration. Subsequent pain assessment will be recorded in the patient diary at 6, 12, 18, 24, 36, 48, and 72 hours and Days 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15. At 24, 48, and 72 hours and Day 5, the assessments will take place just before IMP administration. If the time point for a specific assessment occurs during sleep, pain intensity will be recorded just before and after sleep, and the assessment closest to the scheduled time point will be used in the analysis (see Section 18.16).

For each time point (6, 12, 18, 24, 36, 48, and 72 hours and Days 5, 6, 7 and 8), the change from baseline in pain intensity on VAS is defined as:

VAS pain intensity at specific time point – VAS pain intensity at baseline

18.18 Pain intensity on Likert

In addition to the VAS scale, the patients will score their current pain intensity in the index joint on a 5-point Likert scale (0 to 4 point scale: “none”, “mild”, “moderate”, “severe”, “extreme”).

For each flare, the pain intensity will be assessed in the index joint and recorded in the patient diary prior to the first IMP administration. Subsequent pain assessment will be recorded in the patient diary at 6, 12, 18, 24, 36, 48, and 72 hours and Days 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15. At 24, 48, and 72 hours and Day 5 the assessments will take place before IMP administration. If the time point for a specific assessment occurs during sleep, pain intensity will be recorded just before and after sleep, and the assessment closest to the scheduled time point will be used in the analysis.

For each time point (6, 12, 18, 24, 36, 48, and 72 hours and Days 5, 6, 7 and 8), the change from baseline in pain intensity on Likert is defined as:

Likert pain intensity at specific time point – Likert pain intensity at baseline

18.19 Onset of effect

Onset of effect is defined as ≥20 % change from baseline pain intensity on VAS, and time to onset of effect is defined as the first time point after first IMP injection when this is observed using all measurements at all actual timepoints. The percentage of change from baseline in pain intensity on VAS is defined as:

(VAS pain intensity at specific time point – VAS pain intensity at baseline) / (VAS pain intensity at baseline)

18.20 Response

Response is defined as ≥50 % change from baseline pain intensity on VAS, and time to response is defined as the first time point after first IMP injection when this is observed using all measurements at all actual timepoints. The percentage of change from baseline in pain intensity on VAS is defined as:

(VAS pain intensity at specific time point – VAS pain intensity at baseline) / (VAS pain intensity at baseline)

18.21 Resolution of pain

Resolution of pain is defined as <10 in pain intensity on VAS, and time to resolution of pain is defined as the first time point after first IMP injection when this is observed using all measurements at all actual timepoints.

18.22 Time to rescue medication

Intake of rescue medication is recorded by the patient using the eDiary:

101 = Acetaminophen
102 = Paracetamol
103 = Acetaminophen/Codeine
104 = Paracetamol/Codeine
105 = Codeine
106 = Tramadol
(107 = Ice/cold-pack will not be counted as rescue in the time to rescue calculation)
108 = Other (non medicinal products will not be counted as rescue)

If other is selected the type will be recorded in the Rescue Medication in the eCRF. Patients still having insufficient relief with the rescue medication listed above after the 72 hour assessment are allowed to take glucocorticoids which can be recorded in the Rescue Medication form or Concomitant Medication form in the eCRF. If a patient has taken glucocorticoids, but has no earlier intake of rescue medication from the list, the glucocorticoid intake will be counted as rescue.

The date and time when the first rescue medication is taken in either system will be used in the calculation of time to intake of rescue medication, defined as (in hours):

Date/time of intake of first rescue medication – date/time of intake of first IMP

In case of incomplete or missing date and/or time for either occasion, the following rules will be applied:

- If the minute is missing, the time will be set to the full hour before, i.e. HH:00. If this imputation gives a time point that is prior to the first dose of IMP, the time will be set to the same time point as the first IMP injection.
- If both the hour and minutes are missing, the time will be set to 00:00.
- In the event of missing date, timing of rescue medication cannot be assessed and this recording will not be included in the analysis.
- For glucocorticoids captured as concomitant medication where only date and no timing information is available time is set to 00:00

18.23 Drug exposure

Drug exposure will be calculated as the total amount IMP (mg) given for each flare and for the total study.

For each flare, the patient should receive 11 syringes. At Visit 1 (Day 1), 1 injection of triamcinolone or placebo and 2 injections of anakinra and/or placebo will be given on site and 4 syringes of anakinra and/or placebo will be dispensed for 2 days of treatment at home at 24 and 48 hours. At Visit 2 (Day 4), 2 injections of anakinra and/or placebo will be given on site and 2 syringes of anakinra and/or placebo will be dispensed for 1 day of treatment at home at Day 5. At Visit 2 (Day 4) and Visit 3 (Day 8), unused syringes will be returned.

For patients randomized to anakinra 100 mg, one syringe will contain 100 mg anakinra and one syringe will contain placebo.

For each day, the number of used syringes will be recorded (in the CRF if on site and in the eDiary if at home). For anakinra, the pair of syringes are numbered 1 and 2 and is randomly linked to anakinra or placebo. If only a single syringe is used, its number is recorded.

The number of dispensed syringes for use at home and the number of returned unused syringes are recorded in the CRF dispense log.

Anakinra:

- Total amount given for one flare will be calculated as number of syringes containing Anakinra used*100 mg.
- If a treatment at home is not recorded in the diary but no corresponding unused syringes are returned, it is assumed that the syringes were used (If the number of syringes injected as recorded in the eDiary is less than unused syringes returned recorded in CRF dispense log, the number of syringes used for the flare will be updated to match the dispense log.)
- If a patient, randomized to anakinra 100 mg, has only used a single syringe the number of the syringe will be taken into account, i.e. dependent on if the number is linked to anakinra or placebo, 100 mg or 0 mg will be used in the exposure calculation. If there is no information which syringe was used, it is assumed to be anakinra 100 mg.

Triamcinolone: Total amount for one flare is the same as the exposure at Day 1, and if dose is recorded as administered in the CRF the exposure is 40 mg for that flare.

18.24 Adverse event start/stop day

The start day of an AE will be linked to the specific flare, and is defined as:

Start date of adverse event – Date for 1st IMP injection for latest flare + 1

where the date for 1st IMP injection for latest flare is based on the latest flare prior to start of the specific AE.

The stop day of an AE is defined as:

Stop date of the adverse event – Date for the 1st IMP injection for the latest flare + 1

where the date for the 1st IMP injection for the latest flare is based on the latest flare prior to the start of the specific AE.

For AEs starting within 72 hours after the 1st IMP administration for the latest flare, the start hour will also be calculated and is defined as:

Start date/time hour of adverse event – Date/time hour for 1st IMP injection for latest flare

In addition for the subsequent flares the start and stop dates will also be calculated from the date for the 1st IMP injection for the 1st flare.

In case of incomplete or missing start date, the following rules will be applied:

- If the day is missing, and the month and year is the same as for the date for 1st IMP injection for the latest flare, the start date will be set to the date for 1st IMP injection for the specific flare. Otherwise the date will be set to YYYY-MM-01.
- If day and month are missing, and the year is the same as for the date for the 1st IMP injection for the first flare, the start date will be set to the date for 1st IMP injection for the specific flare. Otherwise the date will be set to YYYY-01-01.
- If day, month and year are missing, the start date will be set to the date for the 1st IMP injection for the first flare, i.e. baseline visit.

In case of incomplete or missing stop date for a resolved AE, the following rules will be applied:

- If the day is missing, the stop date will be set to the last day of the month, i.e. YYYY-MM-30/31 if this date becomes before the end of study date, otherwise the date will be set to end of study date.
- If day and month is missing, the stop day will be set to YYYY-12-31 if this date becomes before the end of study date, otherwise the date will be set to end of study date.
- If day, month and year is missing, the AE will be set to ongoing.

18.25 Treatment emergent abnormal vital signs

Treatment emergent vital signs are defined as: Values that are normal or high at baseline for each flare and shift to low at the corresponding visit, or that are normal or low at baseline and shift to high at the corresponding visit.

The normal reference ranges used are:

- Systolic blood pressure (mmHg): 100 to 140
- Diastolic blood pressure (mmHg): 60 to 90
- Pulse rate (beats/min): 50 to 90

18.26 Immunogenicity

The following definitions based on Shankar et.al. (2014) will apply for this study:

- *Treatment induced ADA*: de novo development following administration of IMP (i.e., formation of ADA any time after the initial IMP administration for the first flare in a patient without pre-existing ADA).
- *Treatment boosted ADA:* Baseline ADA boosted to higher level following IMP administration (i.e., at any time after the initial IMP administration for the first flare the ADA titer is greater than the baseline titer by a factor of four).
- *Treatment-unaffected ADA*: Baseline ADA that does not change following IMP administration.
- *ADA positive patient*: A patient with at least one treatment-induced or treatment-boosted ADA positive sample at any time during the treatment or follow-up observation period.
- *ADA negative patient*: A patient without a treatment induced or treatment-boosted ADA positive sample during the treatment or follow-up observation period.
- *Transient ADA positive patient:* Patient with at least one treatment induced ADA positive sample after baseline and an ADA negative sample at the last sampling time point.
- *Persistent ADA positive patient:* Patient with treatment induced ADA-positive sample at 2 or more sequential sampling time points during the treatment including follow-up period and an ADA positive sample at the last sampling time point.
- *Off-treatment persistent ADA positive patient*: Patient remaining ADA positive after both anakinra drug and any transient ADA would be expected to have cleared (16 weeks after last dose within a flare). ADA detected after this period should be considered persistent off-treatment.

18.27 Short Form (36) Health Survey (SF-36)

The SF-36® Health Survey (acute Version 2) consists of 36 questions that will be scored and grouped into 8 health domains as well as one psychometrically-based PCS score and one MCS score. The calculations of the health domain scores and the PCS and MCS scores will be done

using the QualityMetric Health Outcomes™ Scoring Software 5.0, Version 5.0.6163.22119 using the 2009 US Norm as well as "Maximum Data Recovery" as missing data estimation method.

The 8 health domains are:

- Physical functioning (PF)
- Role-physical (RP)
- Bodily pain (BP)
- General health (GH)
- Vitality (VT)
- Social functioning (SF)
- Role-emotional (RE)
- Mental health (MH)

For each health domain score and PCS and MCS scores, the change from baseline is defined as:

Domain score at specific time point – domain score at baseline

18.28 EuroQol 5 dimensions 5 levels (EQ-5D-5L)

5 dimensions: The EQ-5D-5L consists of 5 dimensions (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) each of which can take one of 5 responses. The responses record 5 levels of severity (no problems/slight problems/moderate problems/severe problems/extreme problems) within a particular EQ-5D dimension, coded as 1 to 5.

For each dimension, the change from baseline is defined as:

Dimension score at specific time point – dimension score at baseline

If the score at either of the time points is missing, no change from baseline will be calculated.

Index value: The descriptive system can be represented as a health state, e.g. health state 21143 represents a patient who indicates slight problems on the mobility dimension, no problems on the self-care and usual activities dimensions, severe pain or discomfort, and moderate problems on the anxiety/depression dimension. These health states will be converted to a single index value using the existing value sets for the US.

VAS: The patients will record their self-rated health on a 0 to 100 VAS. The change from baseline is defined as:

VAS at specific time point – VAS at baseline

If the VAS assessment at either of the time points is missing, no change from baseline will be calculated.

18.29 Work productivity and activity impairment:specific health problems (WPAI:SHP)

The WPAI:SHP contains questions related to the patients quantitative assessment of the amount of absenteeism, presenteeism and daily activity impairment attributable to gout. The questions in the WPAI:SHP records the following:

- Q1 = currently employed
- Q2 = hours missed from work due to specified problem (gout)
- Q3 = hours missed from work due to other reasons
- Q4 = hours actually worked
- Q5 = degree problem affected productivity while working (score 0 to 10)
- Q6 = degree problem affected regular activities (score 0 to 10)

Based on these questions, the percentage of work time missed due to gout, percentage of impairment while working due to gout, percentage of overall work impairment due to gout, percentage of activity impairment due to gout and total work time will be derived. Calculations will only be performed if the patient has responses to all questions involved in a respective formula. If hours worked=0, productivity assessments related to work will not be calculated, only the percentage of activity impairment due to problem.

Currently employed will be set to Yes if Question 1 is answered 'Yes', or if Question 1 is answered 'No' or missing and questions regarding hours missed or worked >0 (Questions 2, 3, and 4).

Percent work time missed due to problem is defined as:

Question 2 / (Question 2 + Question 4)

Percent impairment while working due to problem is defined as:

Question 5 / 10

Percent overall work impairment due to problem is defined as:

[Question 2 / (Question 2 + Question 4)] +

[(1 - (Question 2 / (Question 2 + Question 4))) x (Question 5 / 10)]

Percent activity impairment due to problem is defined as:

Question 6 / 10

Total work time is defined as:

Qustion 2 + Question 3 + Question 4

18.30 Health care resource utilization

At Day 8 and Day 15, patients will be asked about their unscheduled outpatient visits as well as days of hospitalization due to the acute gout flare during the last week. The total number of outpatient visits for Week 1 and Week 2, respectively, is defined as:

No. visits to a Specialist + No. visits to a Nurse + No. visits to a Family Physician

The total number of outpatient visits for a flare is defined as:

Total No. visits for Week 1 + Total No. visits for Week 2

The total number of days of hospitalization for a flare is defined as:

Total No. days of hospitalization for Week 1 + Total No. days of hospitalization for Week 2

19 References

Neogi T, et al. 2015 Gout Classification Criteria. An American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2015;67(10):2557–2568.

European Medicines Agency. Guideline on Missing Data in Confirmatory Clinical Trials. July, 2010.

Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2nd Edition. New York: John Wiley; 2002.

National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials.Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academic Press, 2010.

O'Kelly M, Ratitch B. Clinical trials with missing data: A guide for practitioners. 1st Edition. 2014.

Shankar G, Arkin S, Cocea L, Devanarayan V, Kirshner S, Kromminga A, et al. Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides-harmonized terminology and tactical recommendations. AAPS J. 2014;16(4):658-73.

http://www.reillyassociates.net/WPAI_General.html

Ratitch B, O'Kelly M, Tosiello R. (2013). Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models, Pharmaceut. Statist.2013:12:337-47.

Rubin DB (1976). Inference and missing data. Biometrika. 1976; 63(3):581-92.