Statistical Analysis Plan

Protocol number:	Sobi.ANAKIN-301
NCT number:	03265132
Study title:	A randomized, double-blind, placebo-controlled, multicenter, phase 3 efficacy and safety study of 2 dose levels of subcutaneous anakinra (Kineret®) in patients with Still's disease (SJIA and AOSD)
Version:	Final 1.0
Date:	24 July 2019

Table of contents

1 Abbreviations and definition of terms 5
2 Introduction 6
3 Study objectives and endpoints 7
3.1 Primary objective 7
3.2 Secondary objectives 7
3.3 Study endpoints 8
3.3.1 Primary efficacy endpoint 8
3.3.2 Secondary endpoints supporting the primary objective 8
3.3.3 Key secondary efficacy endpoints 9
3.3.4 Other secondary efficacy endpoints 9
3.3.5 Safety endpoints 10
3.3.6 Pharmacokinetic endpoints 10
3.3.7 Immunogenicity endpoints 10
3.3.8 Productivity endpoints 10
3.3.9 Exploratory endpoints 10
4 Study methods 11
4.1 Overall study design and plan 11
4.2 Selection of study population 12
4.3 Method of treatment assignment and randomization 12
5 Sequence of planned analysis 13
5.1 Interim analyses 13
5.2 Analyses and reporting 13
6 Sample size determination 13
7 Analysis sets 14
7.1 The mITT set 14
7.2 Safety set 14
7.3 Pharmacokinetic (PK) set 14
7.4 Protocol deviations 15
8 General issues for statistical analysis 15
8.1 Time windows 16
8.2 Handling of missing data and outliers 16
8.2.1 Imputation for efficacy endpoints 16
8.2.2 Imputation of Adverse Event start dates 17
8.2.3 Imputation of prior/concomitant medication start/stop date 17

8.2.4 Imputation of date of disease diagnosis 17
8.3 Multicenter studies 18
8.4 Multiple comparisons and multiplicity 18
8.5 Derived and computed variables 18
8.5.1 ACR response 18
8.5.2 Absence of fever 18
8.5.3 JADAS27 19
8.5.4 Inactive disease 19
8.5.5 Primary disease characteristics 19
8.5.6 Treatment exposure 19
8.5.7 Non-compartmental pharmacokinetic analysis 20
9 Patient disposition 21
10 Demographics and baseline characteristics 21
10.1 Demographics 21
10.2 ECG 21
10.3 Primary Still's disease characteristics 21
10.4 Medical history 21
11 Prior and concomitant medication 21
12 Efficacy analyses 22
12.1 Primary efficacy endpoint 22
12.1.1 Sensitivity analyses of the primary endpoint 22
12.2 Secondary endpoints supporting the primary objective 23
12.3 Key secondary efficacy endpoints 23
12.4 Other secondary efficacy endpoints 24
12.4.1 Sustained efficacy in ACR-responders 24
12.4.2 Efficacy of anakinra during 12 weeks treatment 24
12.4.3 Inactive disease 24
12.4.4 Time to study drug discontinuation 24
12.4.5 Glucocorticoid tapering 25
12.5 Subgroup Analyses 25
13 Safety analyses 25
13.1 Treatment exposure 25
13.2 Adverse events 25
13.2.1 Serious adverse events 26
13.2.2 Adverse events leading to withdrawal 26
13.2.3 Deaths 26

13.3 Laboratory data 26
13.4 Vital signs 26
13.5 Physical examination 26
14 Pharmacokinetic analyses 27
15 Immunogenicity analyses 27
16 Productivity analyses 27
17 Exploratory endpoints 27
17.1.1 Population PK/PD analysis 27
17.1.2 Inflammatory biomarkers 27
18 References 28

1 Abbreviations and definition of terms

ACR	American College of Rheumatology
ADA	Anti-drug antibodies
AE	Adverse event
AOSD	Adult-onset Still's disease
ATC	Anatomical therapeutic chemical
AUC_{last}	Area under the serum concentration-time curve from time zero to the last quantifiable concentration (C_{last}), calculated by the linear up-log down trapezoidal method
$AUC_{0\text{-}24h,ss}$	Area under the serum concentration-time curve during a dosage interval
CARRA	The Childhood Arthritis and Rheumatology Research Alliance
CHAQ	Childhood health assessment questionnaire
CL/F	Apparent total clearance of drug from serum after subcutaneous administration
CM	Concomitant medication
C_{max}	Observed maximum serum concentration of anakinra
CRP	C-reactive protein
CSR	Clinical study report
DCRI	Duke Clinical Research Institute
ECG	Electrocardiogram
Hb	Hemoglobin
ICH	International council of harmonization of technical requirements for registration of pharmaceuticals for human use
ILAR	International League of Associations for Rheumatology
IL-1Ra	Interleukin 1 receptor antagonist
IL-6	Interleukin 6
IL-18	Interleukin 18
IMP	Investigational medicinal product
IRS	Interactive response system
mITT	Modified intention-to-treat

JADAS	Juvenile arthritis disease activity score
LLOQ	Lower limit of quantitation
MAS	Macrophage activation syndrome
MedDRA	Medical Dictionary for Regulatory Activities
NAb	Neutralizing antibodies
NCA	Non-compartmental analysis
PD	Pharmacodynamics
PK	Pharmacokinetics
PT	Preferred term
SAE	Serious adverse event
SAP	Statistical analysis plan
SHAQ	Stanford Health Assessment Questionnaire
SJIA	Systemic juvenile idiopathic arthritis
Sobi	Swedish Orphan Biovitrum
SD	Standard deviation
SOC	System organ class
TEAE	Treatment-emergent AE
$t_{½}$	Apparent terminal half-life, calculated by $0.693/\lambda_z$
t_{max}	Time to reach C_{max} following dose injection
VAS	Visual analogue scale
V_d/F	Apparent volume of distribution following subcutaneous administration
WHO	World Health Organization

2 Introduction

This SAP describes the planned analysis and reporting for the Sobi anaSTILLs study (Sobi.ANAKIN-301), a randomized, double-blind, placebo-controlled, multicenter, phase 3 efficacy and safety study of 2 dose levels of subcutaneous anakinra (Kineret®) in patients with Still's disease (SJIA and AOSD).

The purpose of this SAP is to outline the planned analyses to be completed to support the CSR for protocol Sobi.ANAKIN-301. The planned analyses identified in this SAP will be included in

potential 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 CSR.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the FDA and ICH: Guidance on statistical principles in clinical trials (ICHE9). 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 this study is:

- To demonstrate efficacy of anakinra versus placebo in Still's disease as assessed by ACR30 response including absence of fever.

3.2 Secondary objectives

The key secondary objective is:

- To demonstrate early onset of efficacy of anakinra versus placebo in Still's disease.

The secondary efficacy objectives are:

- To evaluate sustained efficacy of anakinra versus placebo in patients that reached at least ACR30 response at Week 2.
- To evaluate efficacy of anakinra versus placebo during 12 weeks treatment.
- To evaluate time to study drug discontinuation in anakinra versus placebo.
- To evaluate glucocorticoid tapering in anakinra and placebo treated patients.
- To evaluate the efficacy of anakinra in the 2 separate dose groups.

The safety objective is:

- To evaluate the safety of anakinra.

The pharmacokinetic objective is:

- To evaluate the PK of anakinra.

The immunogenicity objectives are:

- To evaluate occurrence of ADAs, NAbs and cross-reactivity.
- To evaluate ADAs in relation to safety.
- To evaluate ADAs, and NAbs in relation to efficacy.

The productivity objective is:

- To evaluate absenteeism from school or work.

The exploratory objectives are:

- To explore PK/PD relationship between IL-1Ra/anakinra serum concentrations and selected efficacy and safety parameters.
- To explore the PK properties of anakinra using population analysis.
- To explore the effect of anakinra on the exploratory inflammatory biomarkers (IL-6, IL-18, calprotectin and neopterin) in the treatment of patients with Still´s disease.
- To collect a blood sample for future analysis for genetic factors potentially contributing to the patient's response to anakinra, safety and tolerability. Will be reported separately.

3.3 Study endpoints

3.3.1 Primary efficacy endpoint

The primary efficacy endpoint is:

- ACR30 response at Week 2 with absence of fever attributable to the disease during the 7 days preceding Week 2.

  Definition of ACR30 response: An improvement of ≥ 30% from baseline in at least 3 of any 6 variables listed below. Also, no more than 1 of the 6 variables may worsen by > 30% from baseline.

  1. Physician global assessment of disease activity (VAS)
  2. Patient/parent global assessment of overall well-being (VAS)
  3. Number of joints with active arthritis
  4. Number of joints with limitation of motion
  5. Assessment of physical function (CHAQ/SHAQ)
  6. CRP (mg/L)

  Definition of fever: Body temperature ≥ 38.0 °C (100.4 °F) attributable to the disease

3.3.2 Secondary endpoints supporting the primary objective

The secondary endpoints supporting the primary objective are:

- ACR30 response at Week 1 with absence of fever attributable to the disease during 24 hours preceding Week 1.
- ACR50, ACR70 and ACR90 response at Week 1 and Week 2 with absence of fever attributable to the disease during 24 hours before Week 1 and 7 days preceding Week 2.
- Response in the **individual components** of ACR at Week 1 and Week 2. Response is defined as an improvement of ≥ 30%, 50%, 70% and 90% from baseline.
  - Physician global assessment of disease activity (VAS)

  - Patient/parent global assessment of overall well-being (VAS)
  - Number of joints with active arthritis
  - Number of joints with limitation of motion
  - Assessment of physical function (CHAQ/SHAQ)
  - CRP (mg/L)
- Absence of fever during the 7 days preceding Week 2.

3.3.3 Key secondary efficacy endpoints

The key secondary efficacy endpoints are:

- Absence of fever during the 24 hours preceding Week 1.
- Change from baseline in physician global assessment of disease activity (VAS) at Week 1.
- Change from baseline in patient/parent global assessment of overall well-being (VAS) at Week 1.
- Change from baseline in CRP at Week 1.

3.3.4 Other secondary efficacy endpoints

Other secondary efficacy endpoints are:

- Sustained ACR response at Week 4, Week 8 and Week 12 compared to ACR response at Week 2.
- ACR30, ACR50, ACR70 or ACR90 response with absence of fever 24 hours before Week 1 or during the 7 days preceding the visit at Week 2, Week 4, Week 8 and Week 12.
- Absence of rash 24 hours before Week 1 or during the 7 days preceding Week 2, Week 4, Week 8 and Week 12.
- Change from baseline in CRP, Hb, platelet count and ferritin at Week 1, Week 2, Week 4, Week 8 and Week 12.
- Change from baseline in patient/parent global assessment of disease related pain (VAS) at Week 1, Week 2, Week 4, Week 8 and Week 12.
- Inactive disease at Week 12.
- Change from baseline in JADAS27 at Week 2 and Week 12.
- Time to study drug discontinuation for any reason.
- Time to study drug discontinuation due to lack of efficacy or progressive disease.
- Number of patients who have initiated tapering of glucocorticoids at Week 12.
- Number of patients with decreased dose of glucocorticoids by at least 50% at Week 12 compared to baseline.
- Percentage decrease of glucocorticoid dose at Week 12 compared to baseline.
- Efficacy endpoints as described above, of anakinra in the 2 separate dose groups.

3.3.5 Safety endpoints

The study has the following safety endpoints:

- AEs (including MAS), vital signs and laboratory safety assessments.

3.3.6 Pharmacokinetic endpoints

The study has the following PK endpoints:

- Anakinra trough plasma concentrations and repeated-dose PK parameters at Week 12.

3.3.7 Immunogenicity endpoints

The study has the following immunogenicity endpoints:

- Occurrence of ADAs, NAbs, and cross-reactivity and titer levels of ADA and NAbs at baseline, Week 1, Week 2, Week 4, Week 8 and Week 12.
- Occurrence and titer levels of ADAs in relation to AEs at Week 1, Week 2, Week 4, Week 8 and Week 12.
- Occurrence and titer levels of ADA, including NAb in relation to ACR response and CRP at Week 1, Week 2, Week 4, Week 8 and Week 12.

3.3.8 Productivity endpoints

The study has the following productivity endpoints:

- Number of days off school or work due to Still's disease.

3.3.9 Exploratory endpoints

The study has the following exploratory endpoints:

- Population PK/PD parameter estimates and associated covariates describing intra- and inter-individual variability in respective parameter estimate.
- Population PK parameter estimates and associated covariates describing intra-and inter-individual variability in respective parameter estimate.
- Change from baseline in exploratory inflammatory biomarkers at Week 1, Week 2 and Week 12.

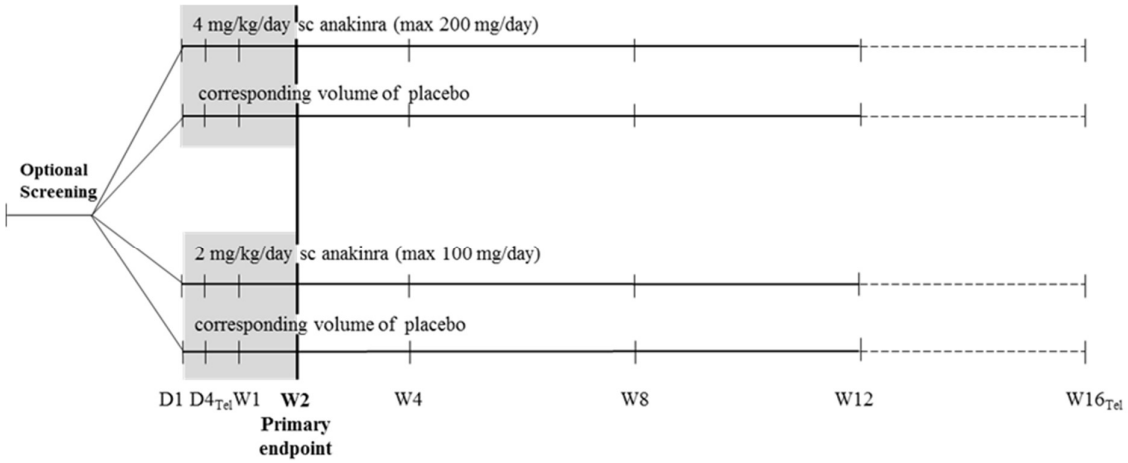
4 Study methods

4.1 Overall study design and plan

The study consists of a 12-week, randomized, double-blind, placebo controlled period with two dose levels of anakinra and a 4-week safety follow-up after last dose of IMP. The primary endpoint will be evaluated at Week 2. Sustained efficacy and time to study drug discontinuation will be evaluated during the full study period.

A screening visit is optional and may be done to identify patients that could be suitable for the study. See Figure 1.

Figure 1 Study design



Abbreviations: D, day, Tel; telephone contact; W, Week.

During the study 6 visits and 2 telephone contacts are scheduled i.e., Day 1 (baseline visit), Day 4_{Tel}, Week 1, Week 2, Week 4, Week 8, Week 12 and Week 16_{Tel} (End of Study).

Patients will be randomly assigned to study drug, after they meet all of the inclusion criteria and none of the exclusion criteria. Patients will receive treatment for 12 weeks, either anakinra or placebo. Patients will be randomized to anakinra in a dose of either 2 or 4 mg/kg/day, with a maximum dose of 100 or 200 mg once daily, respectively. Patients will be randomized to placebo with corresponding volumes for each of the two anakinra dose levels.

To be able to evaluate the primary objective of the study it is important that patients remain in the study until the Week 2 visit whenever possible. However, patients can discontinue study drug at any time during the study e.g. due to lack of efficacy or progressive disease. Patients fulfilling the escape criteria are a subcategory of patients that discontinue study drug due to progressive disease. The escape criteria are symptomatic serositis *or* flare i.e., a worsening of >30 % from baseline in at least 3 of any 6 components part of ACR.

If the investigator judges that a patient must discontinue study drug before the Week 12 visit, a Study Drug Discontinuation visit should be performed before standard of care treatment is initiated, if possible. If the study drug is discontinued on a scheduled visit, the Study Drug Discontinuation visit should be performed instead of the scheduled visit unless the study drug is discontinued at the Week 12 visit. For a patient that discontinues study drug before the Week 2 visit, the Week 2 visit should still be performed according to the original schedule of assessments. However, if the Study Drug Discontinuation visit is performed within 3 days prior to the Week 2 visit, the scheduled Week 2 visit should be omitted. A final follow-up call will be conducted 4 weeks after the last IMP administration according to the assessments described for the Week 16_{Tel} visit in the protocol.

All patients included in the study will be carefully monitored by the investigator and if there is a need, patients can withdraw from the study at any time.

If a patient is on glucocorticoid treatment the dose can be tapered starting earliest at the Week 2 visit. The patient must have reached at least ACR50 response with no fever in the preceding 7 days at the visit initiating tapering.

If a patient develops MAS, the patient will terminate the study drug and will be treated according to standard of care. If the MAS event occurs before the Week 2 visit it is preferred, if possible, that the patient returns to conduct the Week 2 visit according to the schedule of assessments. For all patients that are diagnosed with MAS, irrespective of timepoint in the study, a follow-up call should be conducted 4 weeks after the last IMP administration according to the assessments described for the Week 16_{Tel} visit before the patient is withdrawn from the study. MAS is defined as an event of special interest in this study and must be reported as an SAE.

Once all patients have conducted the Week 16_{Tel} (End of Study) visit or their last visit in case of early discontinuation, data cleaning activities will be completed and the database will be locked before the final analysis will be conducted.

4.2 Selection of study population

A total of 81 patients was originally planned to be randomized, but fewer patients are expected since recruitment will be stopped earlier. At least 12 patients are expected to be randomized, with a 2:1 allocation (anakinra:placebo).

For inclusion into the study, a patient should have active Still's disease, diagnosed not more than 6 months prior to randomization. Diagnosis of Still's disease:

a) If < 16 years of age at disease onset, according to adapted ILAR criteria i.e., CARRA-criteria for SJIA.
b) If ≥16 years of age at disease onset, according to Yamaguchi criteria.

4.3 Method of treatment assignment and randomization

The different treatment groups are;

- placebo (corresponding volume of anakinra 2 mg/kg/day [max 100 mg/day]), or
- placebo (corresponding volume of anakinra 4 mg/kg/day [max 200 mg/day]), or
- anakinra 2 mg/kg/day (max 100 mg/day) or
- anakinra 4 mg/kg/day (max 200 mg/day)

The ratio between the treatment groups is 1:1:2:2, i.e. the ratio placebo:anakinra is 1:2.

The randomization numbers are generated in blocks. Each block includes the four treatment groups per the ratio described above. The block size will not be revealed before breaking of the blind.

The randomization is stratified by age at onset of disease (< 16 years, ≥ 16 years) and glucocorticoid use at inclusion (yes, no). The original plan was to randomize 81 patients with the requirement to randomize at least a third of the patients with an age at onset of disease < 16 years and a third of the patients with an age at onset ≥ 16 years, however this is no longer a requisite due to early stopping of the study. Medidata balance randomization and trial supply system is used for IRS and Biostatistics at DCRI is responsible for generating the randomization scheme, which will link sequential patient randomization numbers to treatment codes.

5 Sequence of planned analysis

5.1 Interim analyses

There is no planned interim analysis for this study.

5.2 Analyses and reporting

All final analyses identified in the SAP will be performed after the last patient has completed the study. The SAP will be finalized, locked and signed prior to database lock and subsequently the breaking of blind.

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

6 Sample size determination

Assuming the ACR30 response rate with absence of fever at Week 2 is 65 % in patients receiving anakinra and 25 % in placebo patients, 81 evaluable patients (54 anakinra and 27 placebo) would be required to ensure 90 % power in demonstrating that anakinra improves clinical features of Still's disease using a two-sided test at a 5 % significance level. These assumptions are based on the canakinumab and tocilizumab phase 3 clinical studies in SJIA where 65 to 85 % of active patients and 10 to 25 % of placebo patients responded at two weeks,

as well as the Nordstrom study in AOSD where 50 % of anakinra treated patients were in remission after 4 weeks of treatment (Nordstrom et al 2012).

Since recruitment will be stopped before the planned number of evaluable patients have been included, it is anticipated that at least 12 patients will be randomized and included in the analyses.

7 Analysis sets

The analysis sets to be used in the statistical analyses are presented in Section 7.1 (mITT set), Section 7.2 (Safety set), and Section 7.3 (PK set).

7.1 The mITT set

The mITT set is the primary analysis population and will comprise all randomized patients, except those who failed to satisfy major disease-specific entry criteria, or who received no IMP (see also section 7.4). The patients will be grouped according to randomization. The mITT is the primary analysis population for the primary and secondary efficacy endpoints.

7.2 Safety set

The Safety set will comprise all patients who received at least one dose of IMP. Patients will be grouped according to the actual treatment received. Patients randomized to placebo that incorrectly received one dose of anakinra or more will be included in the corresponding anakinra group. Patients randomized to anakinra who incorrectly received only placebo will be included in the placebo group.

The Safety set will be used for the safety analyses.

7.3 Pharmacokinetic (PK) set

The PK set will comprise patients who received anakinra without any major protocol deviations as judged by the Clinical Pharmacology Scientist before DBL with respect to administration of IMP. All patients with at least one PK measurement will be included in the analysis set for the analysis of anakinra plasma concentrations. Patients with repeated dose PK should not have any major protocol deviations with respect to the repeated dose PK measurements.

The PK set will be used for evaluating the PK of anakinra.

7.4 Protocol deviations

Protocol deviations related to study inclusion or exclusion criteria, conduct of the study, patient management or patient assessment will be reviewed prior to DBL and described in the CSR as appropriate.

Major protocol deviations that will lead to exclusion from the mITT analysis set are the following:

- Patient failed to satisfy major disease-specific entry criteria in accordance with the adapted ILAR or Yamaguchi criteria, as detailed in the CSP.
- Patient did not get any dose of IMP.

8 General issues for statistical analysis

The comparison of interest is between anakinra (2 mg/kg and 4 mg/kg combined) and placebo. Due to the small number of patients in the individual dose groups of anakinra, summary tables (except for the drug exposure summary) will not provide results by individual dose groups. Likewise, the randomization was stratified by age at onset of disease (< 16 years, ≥ 16 years) and glucocorticoid use (yes, no), but due to the very small number of patients for levels of the stratification factors within each treatment group, descriptive statistics and statistical analyses will generally not be performed by stratification factors. For primary and key secondary efficacy endpoints, dose-level information and stratification allocation will be included in a patient listing.

All statistical tests will be two-sided and performed using a 5 % significance level, if not stated otherwise. Results will be presented as the estimated value for each treatment group, anakinra and placebo, the estimated difference between groups, the associated 95 % two-sided confidence interval and p-value. P-values from statistical analyses will be presented to three decimal places with values below 0.001 displayed as <0.001 and 95 % 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 indicated. Minimum and maximum will be presented to the same number of decimal places as the raw data and mean, SD, 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 9.4 (SAS Institute Inc, Cary, North Carolina, United States).

8.1 Time windows

For the purpose of the statistical analysis and if not otherwise specified, all the variables (primary, secondary and exploratory variables) will be assigned to the visit/assessment timepoint in which they were collected depending on the following analysis time windows in Table 1.

Table 1 Time windows

Visit	Week	Study Day	Study Day window
1	Baseline	1	≤1
2	1	8	2-11
3	2	15	12-22
4	4	29	23-43
5	8	57	44-71
6	12	85	≥72

Unscheduled visit data will be considered when assigning assessments to the analysis visit. If more than one assessment of the variable (scheduled or unscheduled) falls in the same time window but on different days, the closest to the scheduled visit day will be taken, or the earlier, in the event the values are equidistant from the nominal visit date. If several measurements are collected during the selected day, the average (for numeric values)/worst (for categorical values) of the measurements will be taken.

In summaries of extreme values as opposed to visit-based summaries, all post baseline values collected are used including those collected at unscheduled visits regardless if the value is closest to the scheduled visit date or not. For summaries at a patient level, all values should be included, regardless of whether they appear in a corresponding visit-based summary, when deriving a patient level statistic such as a maximum.

Listings will display all values contributing to a time point for a patient.

8.2 Handling of missing data and outliers

8.2.1 Imputation for efficacy endpoints

For all endpoints related to the primary objective the following imputation will be used:

- Patients that discontinue study drug prior to Week 2 will be set to non-responders.
- For patients with missing information in any of the ACR30 components, that specific component will be set to no change. For patients with missing information on fever, the patients will be treated as having a presence of fever.

For the key secondary endpoints the following imputation will be used:

- Patients who have missing information on fever during 24 hours preceding Week 1 will be treated as having a presence of fever in the analysis.
- For the other secondary endpoints no imputation will be performed, i.e., all presentations will be based on observed data.

8.2.2 Imputation of Adverse Event start dates

To assess if AEs are treatment emergent, missing onset dates will be imputed according to the following rules:

- If day is missing and month and year of AE onset is the same as the month and year of first dose of study treatment, then the onset date of the AE will be assumed to be the same as the date of first dose of study treatment.
- If day and month is missing and year of AE onset is the same as year of first dose, then the date of onset will be assumed to be the same as the first dose of study treatment.
- Otherwise, if day is missing, then use YYYY-MM-01, and if day and month are missing, then use YYYY-JAN-01.
- Completely missing onset dates will not be imputed. AEs with a completely missing onset date will be considered treatment-emergent.

8.2.3 Imputation of prior/concomitant medication start/stop date

To assess if medications are prior or concomitant, missing onset dates will be imputed according to the following rules:

- If day is missing and month and year are both available, use 01 of the month or the date of first dose of study treatment whichever is later.
- If only the year is available, use 01 Jan of the year or the date of first dose whichever is later.
- Completely missing onset dates will not be imputed.

Likewise, to assess if a medication is prior, missing stop dates will be imputed according to

- If the day is missing: Assume the last day of the month.
- If the month is missing: Assume 31 Dec of the year.
- Completely missing stop dates will not be imputed

8.2.4 Imputation of date of disease diagnosis

Missing dates for the diagnosis of Still's disease will be imputed according to the following rules:

- If day is missing, 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
- Completely missing date will not be imputed

8.3 Multicenter studies

No effects of centers will be evaluated in this study due to the small numbers at individual sites.

8.4 Multiple comparisons and multiplicity

A fixed sequential testing will be used in the sense that the key secondary efficacy endpoints will only be tested if the null hypothesis related to the primary efficacy endpoint is rejected. No further adjustments will be applied. Testing of all secondary efficacy endpoints will be exploratory.

8.5 Derived and computed variables

8.5.1 ACR response

For efficacy endpoints, ACR30 with absence of fever attributable to the disease, is defined as an improvement of ≥ 30% from baseline in at least 3 of any 6 variables listed below. Also, no more than 1 of the 6 variables may worsen by > 30% from baseline.

1. Physician global assessment of disease activity (VAS)
2. Patient/parent global assessment of overall well-being (VAS)
3. Number of joints with active arthritis
4. Number of joints with limitation of motion
5. Assessment of physical function (CHAQ/SHAQ)
6. CRP (mg/L)

Correspondingly, ACR50, ACR70 and ACR90 with absence of fever attributable to the disease, is defined for ≥ 50%, 70% and 90% response, respectively. Likewise, in these analyses, no more than 1 of the 6 variables may worsen by > 30% from baseline.

Sustained ACR30 response include patients that achieved at least ACR30 response at Week 2, and further had at least ACR30 response at Week 4, Week 8 or Week 12, respectively.

8.5.2 Absence of fever

Separate eCRF forms will specifically query whether fever was absent or present in the 24 hours preceding Week 1 or absent/present in the 7 days preceding Week 2.

Absence of fever at subsequent weeks (Week 4, 8 and 12) will be derived based on information recorded in the fever diary during the 7 days preceding each visit. The number of patients with

no missing data on fever will constitute the denominator for the percentage of patients with absence of fever during each week.

8.5.3 JADAS27

The JADAS27 includes 4 measures: physician global assessment of disease activity, patient or parent global assessment of overall well-being, 27 active joint count, and CRP. The JADAS27 includes the following 27 joints: cervical spine, elbows, wrists, metacarpophalangeal joints (from first to third), proximal interphalangeal joints, hips, knees, and ankles.

JADAS27 is calculated as the sum of its four components, physician global assessment of disease activity converted to cm from the VAS (0=no activity, 10=maximum activity); patient global assessment of well-being converted to cm from the VAS (0=very well, 10=very poor); active joint count (0-27); and CRP. Prior to calculation CRP is truncated to a 0 – 10 scale according to the following formula: (CRP (mg/l) −10)/10. Before calculation, CRP values <10 mg/l are converted to 10 and CRP values >110 mg/l are converted to 110. The JADAS27 tool yields a global score of 0–57.

8.5.4 Inactive disease

At the Week 12 visit and at the Study Drug Discontinuation visit, if applicable, the presence or absence of specific physical examination findings which are associated with Still's disease will be recorded to assess "inactive disease". Inactive disease is a composite of the following parameters: no joints with active arthritis, no fever, no rash, no serositis, no splenomegaly, no generalized lymphadenopathy attributable to Still's disease, CRP level within normal limits, physician's global assessment of disease activity score below 10 mm on a 100 mm VAS and a documented morning stiffness ≤15 minutes. The assessment of inactive disease is based on observed data for all parameters, with no imputation in case of missing information.

8.5.5 Primary disease characteristics

Age at Still's disease diagnosis will be calculated based on date of birth and the date of Still's disease diagnosis as recorded on the Demography eCRF form.

Disease duration will be derived based on the Day 1 visit date and the date of Still's disease diagnosis, as recorded on the Demography eCRF form.

8.5.6 Treatment exposure

The anakinra doses that will be evaluated are 2 mg/kg/day (max 100 mg/day) and 4 mg/kg/day (max 200 mg/day). The dose will be adjusted at baseline according to actual body weight (rounded to the nearest kg) and will remain the same throughout the study. The IMP administration will be supervised by the investigator or investigational site staff at the Day 1 visit and the exact time point for the administration will be recorded. At all other occasions when the

IMP is self-administered the patient will record the exact time point of the administration in the diary.

For each patient, the daily dose will be calculated as the total exposure (amount IMP (mg) given for the entire study) divided by the number of days on treatment.

The total treatment duration will be calculated as (date of last IMP - date of first IMP) +1.

8.5.7 Non-compartmental pharmacokinetic analysis

The serum anakinra concentration data from full PK sampling will be analyzed by NCA using SAS version 9.4. The following PK variables will be derived where applicable:

C_{max}	Observed maximum serum concentration of anakinra
t_{max}	Time to reach Cmax following dose injection
AUC_{last}	Area under the serum concentration-time curve from time zero to the last quantifiable concentration (Clast), calculated by the linear up-log down trapezoidal method
$AUC_{0\text{-}24h,ss}$	Area under the serum concentration-time curve during a dosage interval
V_d/F	Apparent volume of distribution following subcutaneous administration
λz	Terminal slope
$t_{½}$	Apparent terminal half-life, calculated by $0.693/\lambda z$
CL/F	Apparent total clearance of drug from serum after subcutaneous administration

The individual serum concentration data, and the actual time for anakinra administration and blood sampling will be used throughout the analyses. Samples with serum concentrations below the lower limit of quantitation (LLOQ) at early time-points will be treated as zero. Serum concentrations below the LLOQ appearing in the terminal samples will be omitted from the analysis.

Cmax and tmax will be derived directly from observed serum anakinra concentration values. AUClast will be calculated using the linear/log trapezoidal rule.

Extrapolation to 24h will be performed by adding the residual area up to 24h, ($C_{8h,\ pred}/\lambda z$-$C_{24h,\ pred}/\ \lambda z$), where λz will be determined by linear regression of the log terminal part of the serum concentration versus time curve, $C_{8h,\ pred}$ will be the predicted (according to the regression line) concentration at the time of the last sample (8h) and $C_{24h,\ pred}$ will be the predicted (according to the regression line) concentration at 24h. CL/F will be calculated as Dose/$AUC_{0\text{-}24h,ss}$ and V_d/F will be calculated as λz x CL/F.

9 Patient disposition

The number and percentage of patients who were enrolled, randomized, who initiated and who completed or discontinued treatment early (before Week 2 and anytime), and who completed or discontinued the study early (before Week 2 and anytime), together with the reasons for discontinuations, will be presented by treatment group.

The number of patients in the mITT set and in the safety set will be presented.

10 Demographics and baseline characteristics

10.1 Demographics

Baseline characteristics and demographic data, including age, sex, race, and ethnicity will be presented by descriptive statistics for the mITT set and for the Safety set.

10.2 ECG

A baseline 12-lead ECG will be recorded before the first IMP administration, at the Day 1 visit. The number and percentage of patients with abnormal ECG findings will be presented. Any clinically significant abnormalities are also recorded as medical history, and if judged serious, as a SAE.

10.3 Primary Still's disease characteristics

Still's disease characteristics including age at diagnosis, age at symptom onset and disease duration, will be summarized for the mITT and Safety analysis sets.

10.4 Medical history

All medical history data will be coded using Medical Dictionary for Regulatory Activities (MedDRA), version 21.1.

General medical history at baseline will be presented by SOC and PT for the mITT and Safety analysis sets.

11 Prior and concomitant medication

All prior and concomitant medication will be coded using the WHO Drug dictionary.

The preferred term grouped by ATC level 4 will be used for presentation and sorted in descending order of frequency in total (both 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 summarized separately for both the mITT set and the Safety set.

12 Efficacy analyses

12.1 Primary efficacy endpoint

The primary endpoint, ACR30 response at Week 2 with absence of fever attributable to the disease during the 7 days preceding Week 2, will be analyzed using the mITT analysis set and will use no missing data imputation other than that described in 8.2.1. Patients who have discontinued study drug prior to Week 2 will be treated as non-responders in the analysis.

The null and alternative hypotheses with respect to the primary efficacy endpoint are defined as:

H_0: $P_{anakinra} = P_{placebo}$

H_A: $P_{anakinra} \neq P_{placebo}$

where P is the proportion of patients with ACR30 response as defined for the primary endpoint. Fisher's exact test will be used to test the hypothesis at the two-sided significance level of α=0.05. The proportion of patients with ACR30 response in each treatment group and the difference in response rates between anakinra and placebo with the corresponding 95 % exact (Santner-Snell) confidence interval (Santner et al 1980) will be presented.

A patient listing of ACR response and the response for individual components by visit, with information on dose of anakinra, age at onset of disease and glucocorticoid use at inclusion, will be provided.

12.1.1 Sensitivity analyses of the primary endpoint

As a sensitivity analysis for both the primary analysis of ACR30 (+ absence of fever) at Week 2 and the supportive analysis of ACR30 (+ absence of fever) at Week 1 a comparison of treatment groups including the actual response data for patients who have discontinued study drug will be performed using the ACR assessment closest to Week 2 or Week 1 respectively. However, patients with study drug discontinuation due to progressive disease will still be treated as non-responders as in the main analysis. The same exact methods as for the primary analysis will be used.

If the analysis of the primary efficacy endpoint results in a statistically significant outcome a tipping point analysis will be performed. The number of responders that can potentially be observed among patients with a missing ACR30 outcome at Week 2 (e.g. when non-response is assigned in the primary analysis due to discontinuation of study drug) in the anakinra group versus the placebo group will be presented in a basic tipping point display. Pairs of number of responders in the anakinra and placebo group that lead to statistical significance will be marked

and pairs that lead to non-significance will not be marked. The tipping point boundary is the staircase region between marked and unmarked rectangles (Yan et al 2009).

12.2 Secondary endpoints supporting the primary objective

For ACR30, ACR50, ACR70 and ACR90, with absence of fever in the preceding 24 hours at Week 1 and ACR50, ACR70 and ACR90, with absence of fever in the preceding 7 days, at Week 2, the proportion of patients with a response in each treatment group, along with 95 % exact confidence intervals and exact tests of the difference in response between treatment groups, will be presented. The same exact methods as for the primary endpoint will be used. Patients who have discontinued study drug prior to the relevant week of the analysis (Week 1 or Week 2) will be treated as non-responders in the analyses.

The response in the individual components of ACR30, ACR50, ACR70 and ACR90, respectively, at Week 1 and Week 2 will be analyzed in the same way as the primary endpoint using exact statistical methods. The proportion of patients with a response in each group, and exact confidence intervals of the difference will be presented. Patients who have discontinued study drug prior to Week 1 will be treated as non-responders in the analyses of the individual components of ACR at Week 1 and patients who have discontinued study drug prior to Week 2 will be treated as non-responders in the analyses of the individual components of ACR at Week 2.

The numbers of patients with response/no response in the individual components of ACR at Week 1 and Week 2 will also be presented in bar charts.

The absence of fever at Week 2 will be analyzed in the same way as the primary endpoint using exact statistical methods. Patients who have discontinued study drug prior to Week 2 will be treated as having a presence of fever in the analysis.

12.3 Key secondary efficacy endpoints

The absence of fever at Week 1 will be analyzed in the same way as the primary endpoint using exact statistical methods.

For each of the three continuous endpoints; change at Week 1 in CRP, in physician's global assessment of disease activity and in patient/parent global assessment of overall well-being, descriptive statistics for each treatment group will be calculated. For the comparison between anakinra and placebo, exact rank tests based on Wilcoxon scores will be used. The exact 95% confidence intervals based on the Wilcoxon rank statistic for the Hodges-Lehmann estimate of the treatment difference, will be calculated.

12.4 Other secondary efficacy endpoints

12.4.1 Sustained efficacy in ACR-responders

To evaluate sustained efficacy in patients that initially responded to treatment at Week 2 (ACR30, ACR50, ACR70 or ACR90 response criteria), the ACR responses at Week 4, Week 8 and Week 12 will be compared to the response at Week 2. The individual response profiles will be displayed in a Swimmer plot with bars representing the time on study drug and markers for the ACR response achieved at each visit. The purpose is to show onset of ACR response and how it evolves over time in relation to time to treatment discontinuation.

12.4.2 Efficacy of anakinra during 12 weeks treatment

The number and percentage of patients with ACR30, ACR50, ACR70 or ACR90 response with absence of fever (during 24 hours before Week 1 or 7 days preceding Week 2, Week 4, Week 8 and Week 12) at Week 1, Week 2, Week 4, Week 8 and Week 12 will be presented in Swimmer plots (ACR response in relation to time on study drug) as mentioned in section 12.4.1.

The number and percentage of patients with absence of fever and absence of rash (during 24 hours before Week 1 or 7 days preceding Week 2, Week 4, Week 8 and Week 12) at Week 1, Week 2, Week 4, Week 8 and Week 12 will be tabulated.

Individual absolute values and percentage change from baseline values over time at Week 1, Week 2, Week 4, Week 8 and Week 12 for CRP, physician's global assessment of disease activity and change in patient/parent global assessment of overall well-being, will be plotted.

The number and percentage of patients with inactive disease at Week 12 will be presented.

Individual absolute values and percentage change from baseline values of JADAS27 at Week 2 and Week 12 will be plotted.

12.4.3 Inactive disease

The number of patients with inactive disease at Week 12 will be presented.

12.4.4 Time to study drug discontinuation

Kaplan Meier curves will be generated for the anakinra and placebo groups of time to study drug discontinuation due to any cause. Patients completing the study without being prematurely discontinued will be censored at Week 12.

Study drug discontinuation due to lack of efficacy or progressive disease will be illustrated in the Swimmer plots specified in section 12.4.1.

12.4.5 Glucocorticoid tapering

The number and percentage of patients who have initiated tapering of glucocorticoids at Week 12 will be presented.

For patients that initiated tapering the percentage decrease of glucocorticoid dose at Week 2, Week 4, Week 6, Week 8 and Week 12 compared to baseline will be presented with descriptive statistics. The number of patients that have been able to decrease the glucocorticoid dose with at least 50 % from baseline to Week 12 will be presented.

12.5 Subgroup Analyses

Due to the small sample size, no subgroup analyses will be performed.

13 Safety analyses

13.1 Treatment exposure

Exposure to study drug (prescribed daily dose, actual daily dose received and duration of treatment) will be summarized by dose group (2 and 4 mg/kg/day) with descriptive statistics.

13.2 Adverse events

All AEs will be coded using MedDRA version 21.1.

A TEAE is defined as an AE with a start date/time after the first administration of IMP and within 28 days after stopping therapy.

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 study 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 treatment, system organ class and preferred term. Separate tables will present TEAEs classified by relationship to investigational product and TEAEs classified by maximum severity. These tables will be presented by PT in descending order of frequency in the anakinra group. Percentages will be based on the number of patients in the safety analysis set for the specific treatment group.

The number of patients with MAS will be reported.

13.2.1 Serious adverse events

The number and percentage of patients with at least one serious TEAE will be tabulated by SOC and PT. SAEs will also be listed. The listing will include information on treatment group, age, sex, race, start/stop date of event, SOC/PT, causality, severity, action taken, and outcome.

13.2.2 Adverse events leading to withdrawal

The number and percentage of patients with TEAEs leading to study withdrawal will be presented by SOC and PT.

13.2.3 Deaths

Details of any deaths will be listed together with SAEs as detailed in 13.2.1.

13.3 Laboratory data

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

The laboratory safety data will be presented as actual values and change from baseline values over time by descriptive statistics. The number and percentage of patients with low, normal, or high laboratory values at baseline versus subsequent visits will be presented using shift tables.

The values at post-baseline visits versus baseline for each laboratory parameter will be presented in shift plots.

13.4 Vital signs

Vital signs (height (at baseline only), weight, blood pressure and heart rate) will be presented as absolute and change from baseline values over time by descriptive statistics. Individual absolute values for weight, blood pressure and heart rate over time will also be plotted.

13.5 Physical examination

A general physical examination will be assessed at Day 1/Baseline and at all visits until the Week 12 visit, and at the Study drug discontinuation visit if performed. Abnormal findings at Day 1 are reported and summarized under medical history and findings during the study are reported as AEs.

14 Pharmacokinetic analyses

Anakinra trough concentrations will be derived from the serum concentration data, and will be summarized at each visit with descriptive statistics.

Repeated-dose PK at Week 12 will be presented using individual patient concentration-time plots on the semi-log and linear scales. Calculated PK parameters at Week 12 will be listed.

15 Immunogenicity analyses

The number and percentage of patients with ADA and NAb and cross-reactivity at baseline, Week 1, Week 2, Week 4, Week 8 and Week 12 will be presented. ADA titers at baseline, Week 2, Week 4, Week 8 and Week 12 will be summarized by descriptive statistics.

Occurrence and titer levels of ADA in relation to AEs will be reviewed.

Occurrence and titer levels of ADA, including Nab in relation to ACR response and CRP will be presented.

16 Productivity analyses

The number of days off school or work due to Still's disease since last visit will be presented by visit for patients and caregivers separately using descriptive statistics.

17 Exploratory endpoints

17.1.1 Population PK/PD analysis

These analyses are not mandatory and may be performed where appropriate. The analyses will be reported separately.

17.1.2 Inflammatory biomarkers

The absolute values and the changes from baseline in IL-6, IL-18, calprotectin and neopterin at Week 1, Week 2 and Week 12 will be presented by descriptive statistics.

18 References

Nordstrom D. et al. Beneficial effect of interleukin 1 inhibition with anakinra in adult-onset still's disease. An open, randomized, multicenter study. Journal of Rheumatology. 2012;39(10):2008-2011.

Santner, T. J., and Snell, M. K. (1980). Small-Sample Confidence Intervals for and in Contingency Tables. Journal of the American Statistical Association. 1980;75:386–394.

Yan X et al. Missing Data Handling Methods in Medical Device Clinical Trials. Journal of Biopharmaceutical Statistics 2009;19(6):1085-1098.