This Statistical Analysis Plan follows the “Guidelines for the Content of Statistical Analysis Plans in Clinical Trials” published by Gamble et al (December 19, 2017) in JAMA, complying with the ICH E9 guideline.

Instructions/explanations are in red. Example or template text is in black.

The text provided in this template is for guidance and must be modified according to the protocol. Additional details should be added, as necessary.

Delete all red text from the completed document.

Administrative information:

|  |  |
| --- | --- |
| Sponsor name | The sponsor is an individual, company, institution, or organization which takes responsibility for the initiation, management, and/or financing of a clinical trial.  [Nevroklinikken, Oslo University Hospital] |
| Sponsor address | Insert the sponsors address, city and country  [Oslo, Norway] |
| EudraCT number / REC no | The EudraCT number is needed for application to regulatory agency (e.g. SLV). If application to regulatory agency is not applicable, provide the ethics committee number.  [2014-002056-40] |
| Trial title | Insert the protocol title  [A randomized, double-blind, parallel-group study to evaluate the safety and efficacy of switching from innovator infliximab to biosimilar infliximab compared with maintained treatment with innovator infliximab in patients with rheumatoid arthritis, spondyloarthritis, psoriatic arthritis, ulcerative colitis, Crohn’s disease and chronic plaque psoriasis: The NOR-SWITCH study] |
| Trial ID | Insert short title and sponsor identifier  [The NOR-SWITCH study, DIA2014-1] |
| Trial registration number | A trial registration number should be provided which uniquely identifies a clinical trial and its existence on a publicly-accessible registry. The International Committee of Medical Journal Editors (ICMJE) mandates the registration of clinical trials in a primary register of the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) or in ClinicalTrials.gov before recruitment of the first patient as a condition of consideration for publication. In Norway, registration in ClinicalTrials.gov is mandatory. This identifier should be clearly listed in all relevant documentation including the protocol and the SAP.  [NCT02148640] |

SAP and protocol version:

|  |  |
| --- | --- |
| SAP version and date: | Sequentially numbering and dating each SAP version avoids any confusion over which document is the most recent. Transparent tracking of version numbers and amendments facilitates trial conduct, review and oversight. The first final version of a document will be Version 1.0. It is recommended that subsequent final documents will have an increase of “1.0” in the version number (1.0, 2.0, etc.). While the document is under review, subsequent draft versions will increase by “0.1”, e.g., 1.1, 1.2, 1.3, etc. When the revised document is deemed final, the version will increase by “1.0” over the version being revised, e.g. the draft 1.3 will become a final 2.0.  [This SAP is version 2.0, dated 7 May 2018] |
| Protocol version | Referencing the version of the protocol being used is helpful as it links the SAP to the protocol and serves as a reminder that the SAP is not a standalone document and needs to be read in conjunction with the corresponding version of the protocol. This avoids the need for the author to duplicate information from the protocol in the SAP. If there have been protocol amendments after the SAP has been written then the SAP needs to be reviewed against the amendments, and updated where necessary. The information in the SAP revision history may be extended to record that the SAP has been reviewed in light of protocol amendments but no changes were required.  [This document has been written based on information contained in the study protocol version 5.0, dated 13 December 2015] |

SAP revision history:

A clear explanation of the changes made between each version of the SAP is essential, along with a justification for the revision and the date. This is important to maintain transparency. After the first version of the SAP is agreed and signed off, the SAP revision history should include the following information: the previous version number, the SAP section changed, details of the change made along with justification for the revision, and date of revision. A justification for each SAP revision is necessary to document the reasons for changes. This ensures the external validity of the trial as it demonstrates that changes are not being made based on unblinded trial data. From a regulatory perspective when SAP revisions occur after unblinded interim analyses have been conducted the people involved in deciding, writing, or approving the SAP should ideally have no knowledge of unblinded data particularly if the trial will be used for a license application. In other situations it may be sufficient for the justification to document the reason for the change is not based upon comparative data and for the approver to have no knowledge of unblinded data. Any analyses specified after locking of the database and/or un-blinding of the treatment allocation should be denoted “post-hoc specified analyses”.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Protocol version | SAP version | Section number changed | Description and reason for change | Date changed |
| [1.0] | [1.0] | [NA] | [NA] | [21 May 2017] |
| [1.0] | [2.0] | [12.1] | [Time to event analyses changed to interval-censored events] | [12 Dec 2017] |
| [5.0] | [3.0] | [No changes required] | [SAP reviewed against protocol amendments] | [7 May 2018] |

**SIGNATURE PAGE**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **PRINCIPAL/COORDINATING INVESTIGATOR:** |  |  |
|  | Name, title Affiliation |  |  |
|  |  |  |  |
|  | Signature |  | Date (dd/mmm/yyyy) |
|  |  |  |  |
|  | **TRIAL STATISTICIAN:** |  |  |
|  | Name, title Affiliation |  |  |
|  |  |  |  |
|  | Signature |  | Date (dd/mmm/yyyy) |
|  |  |  |  |
|  | **QC STATISTICIAN [if applicable]:** |  |  |
|  | Name, title Affiliation |  |  |
|  |  |  |  |
|  | Signature |  | Date (dd/mmm/yyyy) |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**ABBREVIATIONS**

|  |  |
| --- | --- |
| AE | Adverse Event |
| ATC | Anatomical/Therapeutic/Chemical |
| BMI | Body Mass Index |
| BPM | Beats Per Minute |
| CI | Confidence Interval |
| DMC | Data Monitoring Committee |
| ECG | Electrocardiogram |
| eCRF | electronic Case Report Form |
| EDC | Electronic Data Capture |
| MedDRA | Medical Dictionary for Regulatory Activities |
| SAE | Serious Adverse Event |
| SAS | Statistical Analysis System |
| SD | Standard Deviation |
| SOC | System Organ Class |
|  |  |

**TABLE OF CONTENTS**

[1 Introduction 7](#_Toc518897773)

[1.1 Background and rationale 7](#_Toc518897774)

[1.2 Trial Objectives 7](#_Toc518897775)

[2 Trial Methods 8](#_Toc518897776)

[2.1 Trial Design 8](#_Toc518897777)

[2.2 Randomisation 8](#_Toc518897778)

[2.3 Sample size 8](#_Toc518897779)

[2.4 Statistical Framework 9](#_Toc518897780)

[2.5 Statistical Interim Analyses and Stopping Guidance 10](#_Toc518897781)

[2.6 Timing of Final Analysis 11](#_Toc518897782)

[2.7 Timing of Outcome Assessments 11](#_Toc518897783)

[3 Statistical Principles 12](#_Toc518897784)

[3.1 Confidence Intervals and p-values 12](#_Toc518897785)

[3.2 Adherence and Protocol Deviations 12](#_Toc518897786)

[3.3 Analysis Populations 13](#_Toc518897787)

[4 Trial Population 14](#_Toc518897788)

[4.1 Screening Data, Eligibility and Recruitment 14](#_Toc518897789)

[4.2 Withdrawal/Follow-up 15](#_Toc518897790)

[4.3 Baseline Patient Characteristics 15](#_Toc518897791)

[5 Analysis 16](#_Toc518897792)

[5.1 Outcome Definitions 16](#_Toc518897793)

[5.2 Analysis Methods 18](#_Toc518897794)

[6 Safety Analyses 21](#_Toc518897795)

[6.1 Adverse Events 22](#_Toc518897796)

[6.2 Clinical Laboratory Parameters 22](#_Toc518897797)

[6.3 Vital Signs 22](#_Toc518897798)

[7 Statistical Software 22](#_Toc518897799)

[8 References 23](#_Toc518897800)

[8.1 Literature References 23](#_Toc518897801)

[8.2 Reference to Data Management Plan 23](#_Toc518897802)

[8.3 Reference to the Trial Master File and Statistical Master File 23](#_Toc518897803)

[8.4 Reference to other Standard Operating Procedures or Documents 23](#_Toc518897804)

# Introduction

## Background and rationale

The full rationale for undertaking the trial and trial background are explained in detail in the protocol so only a brief synopsis is necessary within a SAP to avoid duplication of information. The synopsis should include justification for undertaking the trial, why the trial is needed and description of the research question. This item would be regarded as essential if the SAP is to be accessible externally (e.g. published in a journal or on a website) but is optional if the SAP is an internal document only.

[

TNF-inhibitors (TNFi) have significantly improved treatment of spondyloarthritis (SpA), rheumatoid arthritis (RA), psoriatic arthritis (PsA), Crohn’s disease (CD), ulcerative colitis (UC), and chronic plaque psoriasis (Ps) during the last 15 years, but the treatment with the originator infliximab treatment (Remicade®) is expensive. Patients starting biologic treatment can receive less expensive biosimilar CT-P13 (Remsima®) in many countries. However, switching stable patients who are doing well on originator infliximab, has been controversial.

]

## Trial Objectives

The trial objectives reflect the scientific questions to be answered by the trial, defining its rationale and scope. This information may be provided in sufficient detail within the protocol, in which case a reference would be sufficient. If the protocol contains insufficient detail as protocols usually target clinical rather than statistical readers, then additional detail may be required within the SAP. The trial hypotheses should be stated as these provide information on the framework (e.g. superiority, non inferiority) and regions of statistical testing (one or two-sided tests).

### Primary Objective

[

The primary objective of this treatment strategy study is to assess if CT-P13 is non-inferior to innovator infliximab (INX) with regard to disease worsening in patients who have been on stable INX treatment for at least 6 months.

]

### Secondary Objectives

[

The secondary objectives of this study are:

* To assess the safety and immunogenicity of CT-P13 compared to INX in patients who have been on stable INX treatment for at least 6 months
* To compare the efficacy of CT-P13 to INX in patients who have been on stable INX treatment for at least 6 months applying generic and disease-specific outcome measures ]

### Exploratory Objectives (if applicable)

[The exploratory objective of this study is:

* To compare cost-effectiveness of CT-P13 to INX in patients who have been on stable INX treatment for at least 6 months

]

# Trial Methods

## Trial Design

Specify the type of trial design. This can influence many aspects such as methods used, risk of bias, trial conduct, costs, results and interpretation. For example, factorial or adaptive designs can involve more complex methods, analyses, and interpretations than parallel group superiority trials. Although most trials use equal randomisation (i.e. 1:1 for two groups), it is still important to provide the allocation ratio. For drug trials, specifying the phase of the trial (I-IV) may also be relevant.

[

The NOR-SWITCH study is designed as a randomized, double blind, controlled, parallel-group, multi- center, single-country, non-inferiority comparative phase IV study. Treatment allocation is a 1:1 ratio. Patients are randomised to either continued INX or switch to CT-P13 treatment.

]

## Randomisation

Details regarding the randomisation process should be provided within the protocol. Additional detail such as the method of randomisation, e.g. minimisation or stratification, specific information relating to block sizes or specific factor levels used within minimisation or stratification should be stored in a restricted access area. Reference to where this information is stored should be provided in the SAP. This is to protect against predictability of the randomisation sequence by those providing clinical input to the SAP. This allows the statistician executing the SAP to identify stratification factors for use according to ICH E9.

[

Eligible patients are allocated in a 1:1 ratio between INX and CT-P13 treatment, using a computer randomization procedure stratified by diagnosis (RA, SpA, PsA, UC, CD, chronic plaque psoriasis). The randomization is blocked within each stratum.

Details of block size and allocation sequence generation is provided in a separate document unavailable to those who enroll patients or assign treatment.

]

[

The randomisation process is described in full within the clinical trial protocol. Details of the randomisation including the final random allocation list are held securely and unavailable to unauthorized trial personnel.

]

## Sample size

The sample size calculation may be included in full in the SAP or a reference to the sample size calculation in the protocol or other document may be provided. The sample size calculation is an important piece of information for every trial as it determines how many patients are required in the primary analysis to ensure the trial is adequately powered to detect a clinically important difference. The size of that minimum clinically important difference may be used to interpret results. Justification of the sample size should be given including, if appropriate, the expected rate of attrition. All relevant information on which the calculation is based e.g., effect size, power, significance level etc., should be provided with any references to support parameter specifications together with details of any software used. Sufficient detail must be provided to enable another statistician to reproduce the calculation.

[

Sample size calculation has been based on the primary endpoint (proportion of patients experiencing disease worsening during the study period of 52 weeks). Data from the Norwegian NOR-DMARD registry showed a yearly INX treatment failure rate of 18% per year in patients on stable INX for at least one year. Expecting a higher occurrence of disease worsening compared to treatment failure, we assume that 30% of the patients will experience a disease worsening within 52 weeks. A non-inferiority margin of 15% was regarded appropriate based on the PLANETRA study and discussions with the Norwegian regulatory body.

If there is truly no difference between INX and CT-P13 treatment on the proportion of patients with disease worsening after 52 weeks, 394 patients (197 in each arm) are required to be 90% sure that the upper limit of a one-sided 97.5% confidence interval (or equivalently a two-sided 95% confidence interval) will exclude a difference in favour of INX of more than 15%. Different combinations of remission rates and non-inferiority margins are summarized in Table 2 and Table 3.

Because of the non-inferiority design, the primary population will be the per-protocol (PP) population. To adjust for protocol violators (estimated to 20%), a total of 492 (246 in each arm) will be randomized.

]

## Statistical Framework

Specifying the framework of a trial refers to its overall objective to test the superiority, equivalence or non-inferiority of one intervention from another. However, if for example the main objective is to determine equivalence of the primary outcome, secondary outcomes may be intended to demonstrate superiority. The SAP should clearly specify the framework for each outcome or provide a global statement.

### Hypothesis Test

[

This trial is designed to establish the non-inferiority of CT-P13 compared to INX treatment with regard to disease worsening in patients who have been on stable INX treatment for at least 6 months.

* The primary null hypothesis is that CT-P13 is inferior to INX treatment with regards to the proportion of patients with disease worsening during 52 weeks of treatment by a 15% inferiority margin.
* The primary alternative hypothesis is that CT-P13 is non-inferior with regards to the proportion of patients with disease worsening during 52 weeks of treatment by a 15% inferiority margin.

There is only one identified primary analysis in this trial. All other efficacy analyses will be regarded as supportive or exploratory.

]

### Decision Rule

[

This trial is designed to address a single primary outcome. Non-inferiority is claimed if the primary null hypothesis is rejected on the significance level (alpha) of 0.025 (one-sided). That is, if the upper limit of the 95% two-sided confidence interval for the treatment difference is less than 15%.

]

## Statistical Interim Analyses and Stopping Guidance

Information needed to conduct interim analyses should be detailed including statistical methods to be used, who will perform the analyses, what interim analyses will be carried out and when they will be performed e.g. timing and frequency. If interim analyses are not planned then this should be stated for clarity. If details of interim analyses are included in the protocol, or another document e.g. DMC Charter, then depending on the level of detail given the appropriate document could be referenced to avoid duplication. If separate SAPs have been written for interim analyses then these should be referenced.

If analyses are to be performed on the accruing data at multiple time points then methods must be used to control the type 1 error in order to avoid increasing the risk of a false positive result. Various statistical methods have been developed to control this inflated risk such as Heybittle-Peto or O’Brien-Fleming techniques and the chosen approach should be clearly specified, justified and referenced. The DMC Charter could also be referenced, if applicable.

Details should be provided on the guidelines to be used for stopping a trial early. It should be clear whether a statistical method will be considered within the early stopping guideline.

[

There is one interim analysis planned in this trial to detect superiority for early stopping. It will take place when half (135) of the primary outcome events have occurred, and the corresponding data have been entered, verified and validated according to the data management plan. The primary endpoint will be analysed according to methods detailed in the protocol and the statistical analysis plan. The null hypothesis of no difference between the treatment groups in the primary endpoint will be tested two-sidedly against a significance level of 0.01 at the interim analysis. This follows from applying an asymmetric two-sided test according to an O' Brien-Flemming-like spending function on the upper bound.

The Data Monitoring Committee (DMC) will recommend stopping the trial if there is a safety concern, more than 10% of patients experience severe adverse reactions or the interim analysis concludes with superiority of the intervention. There is a separate DMC Charter detailing the procedures for the interim analysis.

]

[

There will be no interim analyses in this trial.

]

## Timing of Final Analysis

Information on the timing of final analyses should be included, if relevant. Information on timing of final analysis should explain whether all outcomes are analysed collectively or whether timing is stratified by length of follow-up required. Details should be provided on whether there are short- term and long-term outcomes and how they will be reported i.e. will all outcomes be analysed collectively or will the short-term outcomes be published earlier and the long-term outcomes reported at a later date.

[

The main analysis is planned when all patients have concluded 52 ± 4 weeks of treatment, all data up to 52 weeks have been entered, verified and validated and the primary database has been locked.

]

## Timing of Outcome Assessments

The time points at which outcomes are measured is helpful information that can be found in the protocol often in table format. The SAP should either refer to the relevant section of the protocol for details or include this information. If outcomes are required to be measured within a particular time window in relation to each planned visit in order to contribute to the analysis then this should also be specified.

[

For all clinically planned measures, visits should occur within a window of the scheduled visit. Visits outside visit window is regarded a protocol deviation. The target day and visits window is defined in the protocol as:

|  |  |  |
| --- | --- | --- |
| **Visit Label** | **Target Day** | **Definition (Day window)** |
| Screening | -1 | Prior to Day 0 |
| V1. Baseline | Day 0 (Randomization) | Day 0 |
| V2 | 91 | Target day ± 14 days |
| V3 | 182 | Target day ± 14 days |
| Last study visit\* | 364 | 323 to 392 |

\*The last study visit is defined as the visit following the last visit with randomised treatment, and where there is a study end statement.

For analysis and tabulation purposes, we define study time points as

|  |  |  |
| --- | --- | --- |
| **Time Point Label** | **Target Day** | **Definition (Day window)** |
| TP1. Baseline | Day 0 (Randomisation) | Information up to randomisation |
| TP2. Week 12 | 91 | Days 1 to 136 |
| TP3. Week 26 | 182 | Days 137 to 173 |
| TP4. Week 52 | 364 | Days 174 to 392 |

If more than one visit fall into the same time point interval, information on all visits will be used in the analyses.

]

# Statistical Principles

## Confidence Intervals and p-values

Where cut-off values are to be used to declare statistical significance then it is important for authors to document the significance level to be used including whether tests will be one- or two-sided. The significance level used for the primary outcome should be consistent with that used in the sample size calculation but secondary outcomes may use different levels.

Authors should pre-define what methods will be used to conduct any adjustment for multiplicity as different methods can lead to different conclusions. The rationale for adjustment and method(s) chosen should be reported. If no adjustment for multiplicity is planned then an explicit statement should be included. Justification for the absence of multiplicity adjustments for secondary outcomes is probably unnecessary unless a claim is to be made on them, however, for the absence of multiple testing corrections on primary outcomes (e.g. different doses) justification should be given. If gatekeeping methods are to be used then authors should state the order of testing.

The confidence intervals (CI) are essential for the interpretation of the statistical analyses reported for any of the primary or secondary outcomes. The level of CI to be reported should be decided at the design stage to avoid bias being introduced by modification based on trial data. The confidence levels used may be consistent across outcomes or vary by primary, secondary, exploratory and safety outcomes and this should be clearly specified.

[

All calculated p-values will be two-sided and compared to a 5% significance level. If a p-value is less than 0.05, the corresponding treatment group difference will be denoted as statistically significant. All efficacy estimates will be presented with two-sided 95% confidence intervals. As there is only one primary null hypothesis to be tested in this trial, there will be no adjustments for multiplicity.

]

## Adherence and Protocol Deviations

### Adherence to Allocated Treatment

Authors should pre-specify their definition of adherence to the intervention. Non-adherence to the intervention can include not completing the intervention, (e.g. not consuming all prescribed drugs or consuming a lower dose than is prescribed). This may be reported to aid generalizability of results or may be linked to an analysis population specification.

[

Compliance is assessed based on the percent of subjects who have taken the scheduled number of pills. It is defined as:

% compliance = (number of pills taken / number of pills supposed to have been taken)\*100%.

The number of pills supposed to have been taken will be calculated as the duration of treatment (end of study medication – start of study medication + 1) multiplied by 4. In this study 2 pills are taken in the morning and 2 pills in the evening.

]

Along with defining adherence to the intervention it is also crucial to describe how adherence to the intervention will be presented and measured. This process avoids any bias being caused by adherence being defined after unblinding of data.

[

The number and % of participants taking more than 75% of the prescribed treatment will be presented in a table for i) randomisation to visit 3 and ii) visit 3 to visit 4. Results will be provided by treatment group.

]

### Protocol Deviations

A protocol deviation is defined as a failure to adhere to the protocol such as the wrong intervention being administered, incorrect data being collected and documented, errors in applying inclusion/exclusion criteria or missed follow-up visits. A protocol deviation should be defined as major or minor. A deviation may be considered a serious breach if it affects efficacy, the safety, physical or mental integrity of the participants in the trial, or the scientific value of the trial. Protocol deviations should be defined prior to unblinding of data to avoid any bias being caused and due consideration given to inclusion of participants within analysis populations. Protocol deviations may be defined in another document and referenced within the SAP.

[

The following are pre-defined major protocol deviations regarded to affect the efficacy of the intervention:

* Entering the trial when the eligibility criteria should have prevented trial entry
* Discontinuation of intervention prior to 52 weeks
* Major change in concomitant immunosuppressing medication
* Use of prohibited rescue medication
* Received or used other intervention than allocated

]

A description should be provided on how protocol deviations will be summarized. Providing details of whether the deviation is major or minor is helpful if sensitivity analyses are to be conducted by removing patients with major deviations to assess impact on overall conclusions or to align with analysis populations. The approach to summarizing the protocol deviation should also be made clear e.g. number and type of protocol deviations by intervention group or listing of all deviations.

[Protocol deviations are classified prior to unblinding of treatment. The number (and percentage) of patients with major and minor protocol deviations will be summarised by treatment group with details of type of deviation provided. The patients that are included in the ITT analysis data set will be used as the denominator to calculate the percentages. No formal statistical testing will be undertaken]

## Analysis Populations

The analysis populations should be specified in advance. This includes how the analysis populations will be defined and which outcomes will be analysed according to each analysis population. It is important to clearly define populations, even if terms are considered standard. For example, there is no consistent definition of ITT and the phrase has different meanings for different authors.

[

The Enrolled set will include all patients who have provided informed consent and have been included into the study data base.

The Full Analysis Set (FAS) will be defined as all patients randomly assigned to a treatment group having received at least one study treatment infusion after randomisation.

The Safety Analysis Set will include all patients having received at least one study treatment infusion after randomisation.

The Per Protocol Analysis Set (PPS) will include all randomised patients meeting the study eligibility criteria and with no major protocol deviations affecting the treatment efficacy.

]

# Trial Population

## Screening Data, Eligibility and Recruitment

If a trial collects screening data then it is important that the data are appropriately presented to describe the representativeness of the trial sample. This information is not only important for the trial but also important for future trials in the area. The process for screening patients e.g. how patients will be screened and what data will be collected, should be fully described within the trial protocol. According to the CONSORT guidelines as a minimum the number of patients who are assessed for eligibility should be provided with this information presented in a flow diagram, however, more detailed tabulations may be provided. The SAP should describe how this data will be summarized and presented.

The trial inclusion and exclusion criteria should be specified in the protocol. Details of how eligibility data will be summarized should be provided. The CONSORT diagram provides details of the number of patients screened followed by a breakdown of how many patients were eligible and how many were excluded due to violating each inclusion/exclusion criteria.

[

The total number of screened patients and reasons for not entering the trial will be summarised and tabulated.

]

Information included within a CONSORT flow diagram displays the progress of all participants through the trial. The CONSORT guidelines say that “you must complete a flow diagram in order to be compliant with the CONSORT 2010 standard.” They provide a CONSORT flow diagram template that can be used and adapted to create a trial specific flow diagram. All necessary information that is displayed in a CONSORT flow diagram should be listed in the SAP so that it is clear where the patient throughput will begin to be summarized and how, specific follow-up time points that will be presented along with information on withdrawals and loss to follow up. Alternatively, a study specific CONSORT flow diagram template can be included in the SAP highlighting the information that will be collected.

[

A CONSORT flow diagram (appendix A) will be used to summarise the number of patients who were:

* assessed for eligibility at screening
* eligible at screening
* ineligible at screening\*
* eligible and randomised
* eligible but not randomised\*
* received the randomised allocation
* did not receive the randomised allocation\*
* lost to follow-up\*
* discontinued the intervention\*
* randomised and included in the primary analysis
* randomised and excluded from the primary analysis\*

\*reasons will be provided.

]

## Withdrawal/Follow-up

In this section, all the possible levels of withdrawal should be listed, which may differ from trial to trial. Participants may withdraw from the intervention but continue with follow-up; withdraw from follow-up but allow data collected to date to be used; withdraw from follow-up and withdraw consent for data collected to date to be used; or be lost to contact/follow-up. Some clarification within the SAP about how each level of withdrawal will be categorized and presented is important.

[

The status of eligible and randomised patients at trial end will be tabulated by treatment group according to

* completed intervention and assessments
* completed assessments but not intervention
* withdrew consent
* lost to follow-up

]

Timing of withdrawals and lost to follow up is important information. This information allows you to see if there are any patterns in lost to follow up or withdrawals between the different time points and intervention groups. Timing of withdrawal from follow-up or lost to follow up data can be presented in a Kaplan-Meier graph, a table or incorporated into a CONSORT flow diagram. For each follow-up time point information on the number of withdrawals and reasons for withdrawal, number included in the analysis and the number died (if applicable) should be provided.

[

Time from randomisation to treatment discontinuation and time from randomisation to withdrawal/lost to follow-up will be presented graphically using the Kaplan-Meier estimator.

]

Patients can withdraw and be lost to follow up for many different reasons e.g. moved home, unable to participate any longer, withdrawn by clinician reasons, death etc. It is useful for the trial team to attempt to ascertain reasons for all withdrawals and loss to follow up. According to ICH E6 ”Although a subject is not obliged to give his/her reason(s) for withdrawing prematurely from a trial, the investigator should make a reasonable effort to ascertain the reason(s), while fully respecting the subject's rights”. Details of how this data will be assessed and presented should be included in the SAP. This information may be presented by intervention arm within a CONSORT flow diagram or in a table.

## Baseline Patient Characteristics

Presentation of baseline characteristics by trial arm is crucial for every trial as it allows the reader to see whether the characteristics are balanced across intervention groups. Information about which baseline characteristics will be summarized in the final report should be specified. Any factors on which the randomisation has been stratified or minimized should be included so that balance across the randomized groups can be demonstrated.

[

The patient demographics and baseline characteristics to be summarised include age in years, gender, symptom duration, weight, height, BMI, education, smoking status, CRP, ESR, use of concomitant immunosuppressive medication and diagnosis specific disease activity measures.

]

It is important to describe how baseline characteristics will be summarized and presented in the final analysis report. Formal statistical comparisons of baseline data by randomized groups are not normally advocated but if such comparisons are planned these should be justified. It is recommended that prognostic baseline characteristics are presented for the analysis population included in the primary analysis of the primary outcome as well as for all randomized participants in order to assess whether attrition has introduced selection bias and/or destroyed the balance after randomisation.

[

Patient demographics and baseline characteristics will be summarised by randomised treatment arm and overall using descriptive statistics (N, mean, standard deviation, median, 25/75 percentiles, minimum, and maximum) for continuous variables, and number and percentages of patients for categorical variables. There will be no statistical analysis of treatment difference. Any clinical important imbalance between the treatment groups will be noted.

]

# Analysis

## Outcome Definitions

List and describe each primary and secondary outcome including details of:

* Specification of outcomes and timings.
* Specific measurement and units (e.g. glucose control hbA1c (mmol/mol or %))
* Any calculation or transformation used to derive the outcome (e.g. change from baseline, quality of life (QoL) score, time to event, logarithm etc).

The SAP should define each outcome explicitly, clearly identifying primary and secondary variables. If multiple primary variables are used then the considerations outlined in ICH E9 (2.2.5 Multiple Primary Variables) should be explored with direction provided on interpretation. Consistency should be ensured with the chosen approach to adjust for multiplicity. If an outcome is recorded at multiple timepoints it needs to be stated, which of these timepoints are required for the specific outcome and analyses. Detailed explanations should be provided, for example for survival outcomes making it clear what the length of survival is (e.g. calculated from the time of randomisation or time of administration of intervention) and censoring information.

The SAP should identify the specific measurement variable and its units if applicable (e.g. overall survival in days) and provide descriptions and details of any data manipulations or derivations to be performed by the statistician. Detail needs to be provided on what data manipulations or derivations will be performed and how they will be carried out. This may be relevant when data collection units may vary, for example HbA1c measurement in % or mmol or quality of life scores. If the calculation of a score is more complex, but a validated algorithm is available, then providing a reference and a link to the algorithm is sufficient. Scoring, including handling of missing data, should follow that proposed by the instrument developers, unless there is good reason to use an alternative technique, which should be described and justified. Sufficient detail needs to be provided in order for the reader to understand how the scores or results are to be calculated for each outcome.

NOTE: There is a bit of confusion regarding the terminology used for outcomes. Outcome, outcome measure, endpoint, variable etc. are used interchangeable. In this document, outcomes are variables either directly measured or derived from measurements during the trial. It can be a single variable, or a variable measured repeatedly. In the latter case, we adopt the position of the CONSORT statement: “When outcomes are assessed at several time points after randomisation, authors should also indicate the pre‐specified time point of primary interest.”

### General Definitions and Derived Variables

List and describe general definitions and derived variables to be used by more than one outcome.

[

#### Body Mass Index

Body Mass Index (BMI) = Body weight in kilograms divided by the square of the height in meters.

#### Harvey-Bradshaw index

The Harvey-Bradshaw index (HBI) was presented in 1980 as a simpler version of the Crohn's disease activity index (CDAI) to quantify the symptoms of Crohn’s disease. It consists of five domains, general well-being (0-4), abdominal pain (0-3), number of liquid soft stools per day, abdominal mass (0-3) and number of predefined complications. The scores of each sub-domain is summed up to compute the HBI.

]

### Primary Outcome Definition

[

The primary outcome is the occurrence of disease worsening during the 52-week trial period. A patient is defined to have experienced a disease worsening if the patient either has an increase in HBI of ≥ 4 points to a minimum HBI of 7 points, or there is a disease worsening consensus between patient and investigator at any visit during the intervention period of 52 weeks. The primary outcome is dichotomous.

]

### Secondary Outcomes Definitions

[

#### Time to disease worsening

Time to disease worsening is defined as the time from baseline until the first visit where a patient either has an increase in HBI of ≥ 4 points to a minimum HBI of 7 points, or there is a disease worsening consensus between patient and investigator. If a patient does not experience a disease worsening during the trial, the observation is censored at the patient’s last visit. This is a time to event outcome.

#### HBI remission

A patient is defined to be in remission if the patient has a HBI of ≤ 4 points. This is a dichotomous outcome. HBI is measured at all clinical visits, but it is the HBI remission status at 52 weeks which is the outcome of primary interest.

#### Change in HBI from baseline

The change in HBI from baseline is defined as HBI at a given visit minus the HBI at baseline. This is regarded as a continuous outcome. Change in HBI at 52 weeks is the outcome of primary interest.

]

### Overview of Outcomes

If the number of outcomes is substantial, it may be sensible to give a table overview.

[

|  |  |  |  |
| --- | --- | --- | --- |
| Level | Outcome | Timeframe | Type |
| Primary | Disease worsening | During trial | Dichotomous |
| Secondary | Time to disease worsening | During trial | Time to event |
|  | HBI remission | 52 weeks | Dichotomous |
|  | Change in HBI | 52 weeks | Continuous |

]

## Analysis Methods

List and describe each primary and secondary outcome including details of what analysis method will be used, and how the treatment effects will be presented. Conclusions can be affected substantially by the analysis method(s) used, therefore it is extremely important to pre-specify the analysis method(s) so there is no possibility of the method being chosen because it gives the most positive results. If transformations are to be applied, then these should be specified along with the rationale for the transformation and the resulting interpretation.

For each outcome, the SAP should specify what analysis method(s) will be used for statistical comparisons and which trial participants will be included in this analysis if applicable. The SAP should also define what summary measures will be reported such as any descriptive statistics to be displayed, what the unit of each effect estimate will be and whether confidence intervals and p- values will be reported. If more than one method is to be used to analyse the primary outcome, e.g. adjusted and unadjusted for covariates, then the primary analysis method should be identified.

Adjustment for covariates

For each analysis, the SAP should specify whether adjustment will be used, and if so, the covariates to be used (including the categories if applicable), and how these will be included in the model (e.g. as fixed effects, random effects etc.). For the primary outcome, it must be clear whether the adjusted or unadjusted analysis is the primary analysis as failing to pre-specify can lead to bias.

### Primary Outcome

#### Primary Analysis

[

The occurrence of disease worsening will be analysed using a logistic regression model with treatment as dichotomous covariate, adjusted for the nominal covariates diagnosis and study site.

]

#### Summary Measures

[

Descriptive statistics will include number and percentage by treatment group. Descriptive statistics will be based on non-imputed data, thus the number of evaluable outcome measurements at the time of primary interest (52 weeks) will also be presented.

The primary effect estimate will be the adjusted risk difference, computed from the logistic regression effect estimate using the delta method. The adjusted relative risk will also be reported together with the p-value of the null-hypothesis test of no treatment difference from the logistic regression. If the 95% confidence interval of the adjusted risk difference does not include 0, the number needed to treat, calculated as the reciprocale of the adjusted risk difference will be reported.

]

#### Assumption Checks and Alternative Analyses

All statistical tests require that a number of assumptions hold in order for the test to be valid and the conclusions drawn from the analysis to be correct. However, a two-stage analysis may lead to bias. If checks on the underlying assumptions are to be performed, then it is important for these to be pre-specified.

Alternative methods if distributional assumptions are violated

Many distributional assumptions can be checked during blind data review and the SAP updated accordingly. However, if assumptions can only be checked once the treatment allocations are known then the SAP should pre-specify the alternative methods to be used if the underlying assumptions do not hold. The approach taken should be considered carefully as bias may be introduced either by choosing the method of analysis based on the results of tests of assumptions or from performing hypothesis tests in which the underlying assumptions are not upheld. Three possible approaches may be considered: i) pre-specify alternative analyses and how the statistician will choose between them in the SAP so that the process is transparent; ii) select a method of analysis that is robust to assumptions, e.g. survival analysis will be carried out using the restricted mean survival time (RMST) method as this does not assume proportional hazards (PH); or iii) state the method of analysis to be used in the SAP and specify that a sensitivity analysis will be performed using an alternative set of assumptions and the results compared.

[

As there are no continuous covariates in the logistic regression, there will be no assumption checks performed for the primary analysis.

]

#### Missing Data

The majority of trials will have some missing data, which can introduce bias depending on the type of “missingness” (e.g. missing completely at random, missing at random, missing not at random). Therefore, it is important that the SAP states how missing data will be handled and reported including details of any statistical methods and their assumptions, to be used to handle missing data. It should explain if there are any plans to impute missing outcome data, including a list of variables that will be used in the imputation process if multiple imputation (MI) is to be used. Using different statistical methods to handle missing data can lead to differing conclusions so it is crucial to pre- specify exactly what methods will be used under what circumstances, and which will be considered the primary analysis. It is highly recommended that sensitivity analyses are conducted to assess the robustness of trial results when using different methods to handle missing data and these should be clearly described in the SAP.

[

For the primary outcome, missing data occur if a patient has not completed the study within the time window of the final visit (52 ± 4 weeks).

The primary method for imputation will be worst case imputation. For the primary outcome this is occurrence of disease worsening.

]

#### Sensitivity Analyses

Include details of any planned sensitivity analyses The SAP should specify whether any sensitivity analyses will be conducted, and what these analyses will be. The SAP should also include details of the analysis population to be used for each sensitivity analysis. The SAP should state whether there is a minimum percentage of missing data required to trigger the need for sensitivity analyses.

[

1. Restricting the primary analysis to the PPS
2. Unadjusted analysis (logistic regression without adjusting for diagnosis and study site)
3. Missing data handling with best case imputation

]

#### Subgroup Analyses

Include details of any planned subgroup analyses. All pre-planned subgroup analyses should clearly specify the baseline characteristics to be considered, the cut-offs for the subgroup categories, the statistical method that will be used and how the results will be presented (e.g. in a forest plot). However while a large number of subgroup analyses should be avoided due to issues with multiplicity this may be appropriate for example when the aim is to demonstrate consistency across subgroups.

[

Subgroup analyses of the primary outcome will be performed by diagnosis. The subgroup analysis will be done by including a treatment-diagnosis interaction term in the logistic regression model, and the resulting treatment effect by diagnosis will be presented using a forest plot.

]

### Dichotomous Secondary Outcomes

Refer to the primary outcome instruction/explanation text and examples.

#### Main Analysis

#### Summary Measures

#### Assumption Checks

#### Missing Data

#### Sensitivity Analyses

#### Subgroup Analyses

### Continuous Secondary Outcomes

#### Main Analysis

#### Summary Measures

#### Assumption Checks

#### Missing Data

#### Sensitivity Analyses

#### Subgroup Analyses

### Time to event secondary outcomes

#### Main Analysis

#### Summary Measures

#### Assumption Checks

#### Missing Data

#### Sensitivity Analyses

#### Subgroup Analyses

### Additional Analyses

Details of any additional statistical analyses required e.g. complier-average causal effect (CACE) analysis.

Any additional analyses to be conducted should be specified with reasons, a description of the additional analysis and how it will be conducted. This may include pre-specified exploratory analyses that are hypothesis generating or confirmatory of issues identified in other trials.

# Safety Analyses

Sufficient detail needs to be provided on summarizing harms e.g. information on severity, expectedness and causality; details of how AE's are coded or categorised; how adverse events (AE’s) data will be analysed, i.e. grade 3/4 only, incidence case analysis, intervention emergent analysis.

Consideration of safety data is key for every clinical trial. It is important that safety data is reviewed and details are provided in the SAP on how safety data will be summarised in the final analysis report including the analysis population to be used. Information may be provided on the severity, causality and expectedness of the adverse event, information on how the adverse events will be coded or categorised and by whom. The method of summarising the adverse event data should be described ensuring that it is clear whether the descriptive summary will use number of events or number of patients and any analyses to be conducted (e.g. will the adverse events be compared descriptively or will formal statistical testing be undertaken.)

[

General safety evaluations will be based on the incidence, intensity, and type of AEs, and clinically significant changes in the patient’s physical examination findings, vital signs, clinical laboratory results and serum drug concentrations. Safety variables will be tabulated and presented for all patients in the safety set.

]

## Adverse Events

[

Adverse events will be coded using MedDRA, version 17.0E. The investigator records the maximum intensity of each AE using the levels mild, moderate and severe. Adverse events with missing intensity will be considered to be severe. For tabulations, only treatment emerging adverse events (TEAEs) will be presented. TEAEs are defined as AEs with a start date on or after date of first randomised treatment. Any AEs prior to treatment will be listed but not tabulated.

The number (%) of subjects with any TEAEs, with 1, 2 or > 3 TEAEs, with treatment related TEAEs, with treatment emerging serious AEs (TESAE) and TEAEs leading to study drug withdrawal will be summarised by treatment group. The number of events and number (%) of subjects with adverse events by system organ class (SOC) and preferred term (PT) will be summarised by treatment group, overall, for severe AEs and for AEs leading to study discontinuation. In addition, a summary table of AEs reported by >= 20% of all patients will be presented by SOC and PT. A detailed patient narrative will be given for any serious adverse event in the clinical study report in addition to listing. Sub-tabulations by diagnosis will also be presented.

]

## Clinical Laboratory Parameters

[

Safety clinical laboratory parameters were collected and assessed, but only used to identify adverse events. Clinical laboratory parameters will be summarised by treatment group and time point.

]

## Vital Signs

[

Changes in vital signs (including systolic and diastolic blood pressure [mmHg], heart rate [beats per minute] and weight [kg]) will be summarised by treatment group and time point.

]

# Statistical Software

Details of the statistical packages to be used to conduct the statistical analyses may be provided in the SAP. Version numbers of software may change during the time of the trial and should not be specified in the SAP but included in the final reports.

[

All statistical analyses will be done in Stata v14 (StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX, USA).

]

# References

## Literature References

References should be provided in a SAP for any non-standard statistical methods that will be used. If there is any doubt on whether a method is non-standard then it is better to include a reference.

[

1. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. 2017;318(23):2337-2343. doi:10.1001/jama.2017.18556.

]

## Reference to Data Handling Plan

Reference should be made to the Data Handling Plan (DHP) with the version number that was used when writing the SAP. This is important as both documents should be linked with information in the DHP that is also important for the final analysis report. If there is no DMP, then the location of this information (e.g. data handling and cleaning) should be provided.

## Reference to the Trial Master File and Statistical Documentation

The Statistical Documentation contain also copies of the TMF documentation, and is held separately with restricted access . The Statistical Documentation may hold details of the randomisation process or specific protocol deviations that the statistician needs to refer to when executing the statistical analysis plan. If the Statistical Documentation is held separately to the Trial Master File then both should be referenced.

As agreed with the PL, the Statistical Documentation will be kept for <time> for future reference in CTU.

## Reference to other Standard Operating Procedures or Documents

Reference should be made to any other Standard Operating Procedures (SOPs) or documents that are adhered to and followed when writing the SAP.