# Supplement 1

# Protocol and statistical analysis plan

1. Final protocol
2. Summary of protocol changes
3. Original statistical analysis plan, which is also the final version as no updates have been made.

# REmission in rheumatoid arthritis – assessing WIthdrawal of disease-modifying antirheumatic drugs in a Non-inferiority Design

## The ARCTIC REWIND study

**Protocol Identification Number:** DIA2012-1 / ver 4_1
**EudraCT Number:** 2012-005275-14


| | |
|---|---|
| **SPONSOR** | **Diakonhjemmet Hospital AS**<br>Contact person:<br>Professor Kåre Birger Hagen, PhD<br>Box 23 Vinderen, 0319 Oslo<br>Tel : 22 45 15 00<br>E-mail: k.b.hagen@medisin.uio.no |
| **PRINCIPAL INVESTIGATOR** | **Professor Espen A. Haavardsholm, MD PhD**<br>Dept. of Rheumatology<br>Diakonhjemmet Hospital, Box 23 Vinderen, 0319 Oslo<br>Tel: 99798329/22451500<br>E-mail: e.a.haavardsholm@medisin.uio.no |


**PROTOCOL VERSION NO 4_1 - 01.12.2017**

## Study organization

| | |
|---|---|
| **Protocol number / version** | DIA2012-1 / ver 4_1 |
| **EUDRACT number** | 2012-005275-14 |
| **Phase** | 4 |
| **Study design** | Multicenter, randomized, open-label |
| **Date** | December 2017 |

| | |
|---|---|
| **Principal investigator** | Professor Espen A. Haavardsholm, MD PhD |
| **National Medical Leader** | Professor Tore K. Kvien, MD PhD |
| **Biostatistician** | Inge Christoffer Olsen, PhD<br>Joseph Sexton, PhD |
| **Project leader** | Siri Lillegraven, MD MPH PhD |
| **Methodology advisors** | Professor Désirée van der Heijde, MD PhD<br>Professor Daniel H. Solomon, MD MPH<br>Professor Robert Landewé, MD PhD<br>Professor Ivar Sønbø Kristiansen, MD MPH PhD |

| | |
|---|---|
| **Monitors** | Department of Clinical Research Support, Oslo University Hospital *(South-Eastern Norway)*<br>Innovest AS *(Bergen)*<br>Department of Quality, University Hospital of North Norway *(Tromsø)*<br>Department of Research and Development, Helse Møre og Romsdal *(Ålesund)* |

| | |
|---|---|
| **Sponsor** | Diakonhjemmet Hospital AS<br>Contact person: Professor Kåre Birger Hagen<br>Box 23 Vinderen, 0319 Oslo, Norway |

| | |
|---|---|
| **Participating centers** | Diakonhjemmet Hospital<br>Universitetssykehuset i Nord-Norge<br>Helse Møre og Romsdal HF, Ålesund sjukehus<br>Haukeland Universitetssykehus Bergen<br>Sørlandet Sykehus Kristiansand HF<br>Haugesund Sanitetsforenings Revmatismesykehus AS<br>Martina Hansens Hospital<br>Sykehuset Østfold HF Moss<br>Lillehammer Revmatismesykehus<br>Vestre Viken HF Drammen Sykehus<br>Privatpraksis Bendvold / Dovland Kristiansand<br>Revma Vestfold<br>Helgelandssykehuset Mo i Rana |

## Sponsor signature page

| | |
|---|---|
| **Title** | REmission in rheumatoid arthritis – assessing WIthdrawal of disease-modifying antirheumatic drugs in a Non-Inferiority Design |
| **Protocol ID number** | ARCTIC REWIND / DIA2012-1 / ver 4.1 |
| **EudraCT number** | 2012-005275-14 |

**Sponsor signatory approval**
*I hereby declare that I will conduct the study in compliance with the Protocol, ICH GCP and the applicable regulatory requirements:*

Diakonhjemmet Hospital
Professor Kåre Birger Hagen, PhD

Kåre Birger Hagen — 14/12-2017

*Sponsor signature* *Date*

**Coordinating PI signatory approval**
*I hereby declare that I will conduct the study in compliance with the Protocol, ICH GCP and the applicable regulatory requirements:*

Professor Espen A. Haavardsholm, MD PhD

14/12-2017

*PI signature* *Date*

**Local PI signatory approval (Center:_____)**
*I hereby declare that I will conduct the study in compliance with the Protocol, ICH GCP and the applicable regulatory requirements:*

*PI signature* *Date*

## Protocol synopsis

**RE**mission in rheumatoid arthritis – assessing **WI**thdrawal of disease-modifying antirheumatic drugs in a **N**on-inferiority **D**esign

**The ARCTIC REWIND study**

| | |
|---|---|
| **Investigational Product** | Patients receiving tumor necrosis factor inhibitors (TNFi): Step-down and withdrawal of TNFi is compared to continued full-dose TNFi<br>Patients receiving synthetic disease modifying anitrheumatic (DMARD) treatment: Half-dose synthetic DMARD therapy is compared to continued full-dose synthetic DMARD therapy. Patients in the half-dose group are randomized to either stable half dose DMARD therapy or withdrawal after one year. |
| **Centers** | Diakonhjemmet Hospital<br>Universitetssykehuset i Nord-Norge<br>Helse Møre og Romsdal HF, Ålesund sjukehus<br>Haukeland Universitetssykehus Bergen<br>Sørlandet Sykehus Kristiansand HF<br>Haugesund Sanitetsforenings Revmatismesykehus AS<br>Martina Hansens Hospital<br>Sykehuset Østfold HF Moss<br>Lillehammer Revmatismesykehus<br>Vestre Viken HF Drammen Sykehus<br>Revma Vestfold<br>Helgelandssykehuset, Mo i Rana<br>Privatpraksis Bendvold/Dovland Kristiansand |
| **Study Period** | Estimated date of first patient enrolled: March 1st 2013<br>Anticipated recruitment period: March 2013 1st - June 30th 2018<br>Estimated date of last patient completed: June 30th 2021 |
| **Treatment Duration** | 36 months |
| **Follow-up** | 36 months |
| **Main objective** | To assess the effect of tapering and withdrawal of DMARDs on disease activity in RA patients in sustained remission. |
| **Endpoints** | Primary endpoint: Proportion of patients who are non-failures (have not experienced a flare) at 12 months<br><br>Secondary endpoints: Composite disease activity scores and remission criteria, joint damage and inflammation assessed by various imaging modalities, work participation, health care resource use, health related quality of life. |

| | |
|---|---|
| **Study Design** | A randomized, open, controlled, parallel-group, multicenter, phase IV, non-inferiority strategy study with three groups:<br><br>Group B: Patients in sustained remission on TNFi are randomized 1:1 to either stable TNFi treatment OR 4 months of half dose TNFi followed by withdrawal of TNFi. Any co-medication with synthetic DMARDs is kept stable throughout the study.<br><br>Group S: Patients in sustained remission on synthetic DMARDs are randomized 1:1 to either stable synthetic DMARD treatment OR 12 months of half dose synthetic DMARDs. After 12 months, patients in the half dose arm who have not experienced a flare in disease activity are randomized 1:1 to continued stable therapy OR withdrawal of DMARDs.<br><br>ARCTIC follow-up group: Patients from the ARCTIC trial with active disease or less than one year of remission are followed longitudinally. |
| **Main Inclusion Criteria** | - Rheumatoid arthritis (RA) according to specified classification criteria<br>- Group B (patients on TNFi): Any disease duration<br>- Group S (synthetic DMARDs): RA diagnosis after 01.01.2010.<br>- Sustained DAS or DAS28 remission for the last 12 months, DAS remission and no swollen joints at inclusion **OR** participation in the first ARCTIC study |
| **Number of patients** | 360 patients |
| **Efficacy** | Efficacy variables include flare assessment, individual parameters of the ACR core data set, ACR response, DAS and other composite disease activity measures, EULAR response, the van der Heijde modified Sharp score of radiographs, 32 joint PD and grey scale ultrasonography score and MRI scores (RAMRIS, Haavardsholm tenosynovitis score and the OMERACT MRI JSN score). |
| **Safety** | Physical examination and vital signs, laboratory tests, record of adverse events and serious adverse events, linkage to registries for long-term safety outcomes. |

## Table of Contents

## List of abbreviations and definitions of terms

| | |
|---|---|
| ACR | American College of Rheumatology |
| AE | Adverse event |
| ALT | Alanine aminotransferase |
| Anti-CCP | Anti-Cyclic Citrullinated Peptide antibodies |
| AST | Aspartate aminotransferase |
| CDAI | Clinical Disease Activity Index |
| CI | Confidence Interval |
| CR | Conventional Radiography |
| CRP | C-Reactive Protein |
| CSCT | Central Study Coordinating Team |
| CT | Computer Tomography |
| DAS | Disease Activity Score |
| DIP | Distal Interphalangeal |
| DMARD | Disease Modifying Anti-Rheumatic Drug |
| eCRF | Electronic Case Report Form |
| ESR | Erythrocyte Sedimentation Rate |
| et al. | And others |
| EULAR | European League Against Rheumatism |
| FOV | Field Of View |
| GCP | Good Clinical Practice |
| GH | General Health |
| GI | Gastrointestinal |
| HAQ | Health Assessment Questionnaire |
| HLA | Human Leukocyte Antigen |
| ICH | International Conference on Harmonization |
| IEC | Independent Ethics Committee |
| ILAR | International League of Associations for Rheumatology |
| IM | Intramuscular |
| ITT | Intent-to-treat |
| JSN | Joint Space Narrowing |
| LEF | Leflunomide |
| MCP | Metacarpophalangeal |
| MRI | Magnetic Resonance Imaging |
| MTP | Metatarsophalangeal |

| | |
|---|---|
| MTX | Methotrexate |
| NSAID | Non-Steroidal Anti-Inflammatory Drug |
| OMERACT | Outcome Measures in Rheumatology |
| RA | Rheumatoid Arthritis |
| RAMRIS | Rheumatoid Arthritis Magnetic Resonance Imaging Score |
| RF | Rheumatoid Factor |
| PIP | Proximal Interphalangeal |
| QoL | Quality of Life |
| SAE | Serious Adverse Event |
| SDAI | Simplified Diseases Activity Index |
| SD | Standard Deviation |
| SDD | Smallest Detectable Difference |
| SJC | Swollen Joint Count |
| SPSS | Statistical Package for the Social Sciences |
| SRM | Standardized Response Mean |
| STIR | Short Tau Inversion Recovery |
| TB | Tuberculosis |
| TJC | Tender Joint Count |
| TNF | Tumor Necrosis Factor |
| ULN | Upper Limit of Normal |
| US | Ultrasonography |
| VAS | Visual Analogue Scale |
| vdHSS | van der Heijde modified Sharp Score |
| WBC | White Blood Cell |
| WHO | World Health Organization |

# 1 Introduction

This protocol outlines a randomized, multi-center, parallel-group, phase IV, multi-regional non-inferiority strategy trial to study reduction and withdrawal of disease-modifying antirheumatic drugs (DMARDs) in patients in sustained rheumatoid arthritis (RA) remission – the ARCTIC REWIND study.

## *1.1 Background*

RA is a chronic inflammatory disease predominantly affecting women. The main feature is joint inflammation with subsequent joint destruction, but the disease can also have extra-articular manifestations and is associated with increased risk of cardiovascular disease and reduced life-expectancy(1). RA is primarily treated with DMARDs, which are classified as synthetic DMARDs (e.g. methotrexate, sulfasalazine) or biologic DMARDs (e.g. TNF inhibitors, rituximab) (1;2). The implementation of new therapeutics, treat to target principles and modern treatment strategies in RA care has led a significant proportion of patients to reach clinical remission.

### 1.1.1 Remission in RA

The word "remission" implies absence of disease (no signs and symptoms of active disease), but is not the same as "cure", which implies that the disease process will not return. The introduction of biologic DMARDs has changed the outcome of RA, and the goal in modern RA treatment has become to reach and sustain remission, with prevention of structural joint damage and disability (2-5). In addition to the introduction of biologic DMARDs, more aggressive treatment strategies (as exemplified in the first ARCTIC trial) means that patients normally will start treatment with effective medication during the "window of opportunity" – the first months of disease when treatment can improve the long-term outcome (2;6).

The first definition of remission in RA was published as early as 1948 (7), and the first formal definition of remission in RA was developed by an American Rheumatism Association (ARA, later ACR) subcommittee and published in 1981 (8). Due to their stringency, the requirement of two months duration and inclusion of elements not part of the ACR core set for RA, the 1981 ARA criteria have not been widely applied in clinical trials or clinical practice. In the last decade, cut-offs for disease activity indices have been more commonly used to define remission, e.g. the disease activity score (DAS) (<1.6) (9), DAS with 28 joints (DAS28) (<2.6) (10), simplified disease activity index (SDAI) (≤3.3) and clinical disease activity index (CDAI) (≤2.8) (11). In recent years, the need for a consensus on how to define remission led ACR

and EULAR to collaborate to re-evaluate the concept of remission with the aim to reach a new, stringent definition of remission (12-15). Through an elaborate process, including testing of predictive validity with regards to radiographic progression and deterioration of physical function, several candidate remission definitions were assessed. Two definitions of remission were agreed upon: 1) A Boolean-based definition: Tender joint count, swollen joint count, C-reactive protein (CRP, in mg/dL) and patient global assessment (0-10 scale) all ≤1, and 2) An index-based definition: SDAI ≤3.3 (13;14).

The definition of remission is still debated even after the introduction of the ACR/EULAR criteria. Several studies indicate that modern imaging yields information that can help predict disease flare and progressive joint damage in RA remission (16-22), and it has been discussed whether imaging should be implemented in remission criteria (23-26). Other issues commonly discussed are whether examination of the feet is necessary to define remission, and if remission has to been sustained at more than one time point.

### 1.1.2 Reduction of DMARD treatment in RA remission and low disease activity

Some observational studies and clinical trials have assessed reduction of DMARD treatment, but the data for patients in sustained RA remission is limited.

Data on infliximab discontinuation were presented in a post-hoc analysis from the Dutch BeSt study (27). Of patients with six months of low disease activity, infliximab could successfully (defined as sustained low disease activity) be discontinued in 52% (27). It is important to note that these patients did not have to be in remission before infliximab withdrawal. Similarly, 55% of Japanese patients with 24 weeks of low disease activity maintained low disease activity over the next year (28). A small study of patients with long disease duration and sustained remission described 5 of 20 patients as being in remission one year after complete TNF inhibitor withdrawal (29).

In the OPTIMA trial, patients who at 26 weeks had reached low disease activity on combination therapy with methotrexate and adalimumab were randomized to continued treatment with adalimumab and methotrexate or placebo and methotrexate. Most patients who were switched to a combination of placebo and methotrexate maintained good clinical, functional and radiographic through week 78 of the study (30). The HIT HARD study compared induction therapy with a combination of adalimumab and methotrexate to methotrexate and placebo. After 24 weeks, all patients were switched to methotrexate

monotherapy. No significant differences between the groups were found for the primary outcome, disease activity measured by DAS28 at 48 weeks (31).

Two studies indicate that reduced dose of the tumor necrosis factor alpha inhibitor (TNFi) etanercept (ETN) is an option in RA low disease activity. Neither of the studies has been published yet, but both have been presented at international rheumatology meetings. The studies, DOSERA (73 patients, established disease (32)) and PRESERVE (607 patients, early disease), require patients to be in a period of low disease activity before randomization to either full dose ETN + MTX, half dose ETN + MTX, or placebo + MTX. Both studies found comparable results for the ETN + MTX groups, and a significantly worse outcome in terms of disease activity in the placebo + MTX group. However, the populations included in the studies (patients in low disease activity and not remission) somewhat limits the conclusions that can be drawn regarding withdrawal of biologic treatment.

A case series of 15 patients described dose reduction of methotrexate in RA patients in remission through increase in treatment intervals from one week to two weeks. Two patients experienced flares during the 1-year follow-up period, for the other patients disease activity was stable (33).

### 1.1.3 Drug free remission

Drug free remission has been a concept in RA since the seventies, although some early publications assessed withdrawal of synthetic DMARD regimens no longer commonly used (34-37). A randomized controlled clinical trial published in 1996, comparing continued synthetic DMARD treatment to placebo, found flare rates of 22% in the continued therapy group and 38% in the placebo group (36). Although the flare rate was higher in the placebo group, the findings describe 62% of the patients as being in sustained remission on placebo. A weakness of this study is that patients used a multitude of different synthetic DMARDs, none received biologic DMARDs and only a few patients received methotrexate (36). In the more recent BeSt study, 115 of 508 (23%) RA patients achieved drug free remission during the first five years of follow-up, although 46% of these later had to restart DMARDs due to a rise in disease activity ≥ the DAS remission cut-off of 1.6 (38). A comparison of BeSt and the observational Leiden Early Arthritis Cohort showed similar rates of sustained DMARD free remission, and common predictors of this state in both cohorts were absence of ACPA and short symptom duration (39). Treatment strategy was not an independent predictor of drug free remission in these studies. In the Finnish Rheumatoid Arthritis Combination therapy (Fin-RACo) study 14-19% of the patients discontinued DMARDs due to remission (40).

Studies of drug-free remission in RA have been discussed and summarized in several reviews and opinion pieces (27;41-44).

### 1.1.4 Immunogenicity

Biologic therapeutics can be recognized by the human immune system as 'non self' and induce an immune response, also known as immunogenicity. Most therapeutic antibodies induce an unwanted immune response, and the immunogenicity of the first generations of therapeutic antibodies has been quite extensively studied. About 50% of patients treated with chimeric antibodies (e.g. infliximab) develop human anti-chimeric antibodies (HACA) (45-47). Humanization of the variable regions reduces immunogenicity, but even in fully humanized antibodies (e.g. adalimumab) anti-human antibodies (HAHA) will develop in about 20% of treated patients. However, the majority of patients who develop anti-drug antibodies do not make sufficient quantities to neutralize the drug. The presence of either HACA or HAHA is associated with adverse events, of which infusion or injection reactions are the most common. Anti-drug antibodies can also lead to decreased drug levels and consequently impaired treatment responses. These antibody responses may be influenced by a number of patient-related factors such as genetic background, co-morbidity, other immunomodulating therapies (co-medication) and dosing schedule. It has been shown that for example concomitant methotrexate reduces the effect of immunogenicity (48).

Immunogenicity may be a factor when the dosage of medication is increased due to flare. In the case of non-response to previous therapeutic dosage, this may be due to immunogenicity.

## *1.2 The ARCTIC trial*

The ARCTIC trial is a Norwegian randomized, open, prospective, multi-center, parallel-group clinical study of 2 years duration. The main objective is to assess what the treatment target should be to achieve optimal patient outcomes in early RA: Clinical remission or imaging remission? The study is designed with a strategic treatment decision component to evaluate the efficacy and added value of applying ultrasonography compared to standard clinical assessment without the use of information from ultrasonography in patients with early rheumatoid arthritis. The main outcomes are clinical (DAS) remission and inhibition of radiographic progression.

Treatment adjustments in ARCTIC must be made according to a pre-specified algorithm, based on DAS status and change. In the ultrasonography arm the ultrasonography findings can overrule the DAS-derived treatment decision. More than 200 of the planned 240 patients have been included into ARCTIC so far and the last patient is expected to be included in February or March 2013. As the treatment target in ARCTIC is remission and the study applies tight control and "treat to target" principles, it is anticipated that the majority of the patients will be in remission at the end of the study.

### *1.3 Rationale for the study*

#### 1.3.1 Clinical implications

During the last decade, treatment strategies in RA have become much more intensive, aiming for early, aggressive treatment and remission (2;49). This results in a relatively rapid escalation of treatment, with a substantial proportion of RA patients receiving combinations of synthetic DMARDs and/or biologic DMARDs in early disease. DMARDs can potentially cause adverse events, and patients generally prefer to receive as few drugs as possible. The results from ARCTIC REWIND will inform treatment decisions in RA remission and might allow more patients to step down their treatment.

#### 1.3.2 Societal and public health implications

Approximately 25-30 000 individuals in Norway have RA, making it the most prevalent inflammatory joint disease. The total cost of RA to the Norwegian society has been estimated to 6.2 billion NOK, and a recent publication from the Norwegian NOR-DMARD register indicates 2-year costs for RA patients receiving synthetic DMARD treatment of €64 300, with corresponding numbers for patients receiving biologic DMARDs of €121 900 (50). Of prescription-only medicines (POMs), the TNF inhibitors etanercept, infliximab and adalimumab were the top three drugs with regards to sales in 2011 (440, 406 and 390 million NOK respectively (46)). If withdrawal or tapering of TNFi is non-inferior to stable TNFi treatment in deep RA remission, the potential savings to the society will be significant and can fund other important health care measures. In addition, data from synthetic DMARD group in the study will indicate whether drug free remission is realistic in RA. The results of this study will be of international importance.

## 2. Objectives

### *2.1 Primary objective*

To assess the effect of tapering and withdrawal of DMARDs on disease activity in RA patients in sustained remission.

### *2.2 Secondary objectives*

a) To study predictors of successful DMARD reduction and discontinuation in RA remission
b) To evaluate the cost-effectiveness of alternative treatment strategies in stable RA remission
c) To assess whether the choice of DMARD strategy in RA remission influences the level of Magnetic Resonance Imaging (MRI) and ultrasonography detected inflammation
d) To assess joint damage in patients in sustained RA remission who continue to receive stable DMARD treatment and patients who receive reduced DMARD treatment
e) To provide long-term follow-up data on patients included in the original ARCTIC trial
f) To assess relationships between treatment, inflammation, physical function and joint damage in RA remission
g) To examine adverse events rates in different DMARD strategies in RA remission
h) To examine the value of imaging information in treatment decision making in sustained RA remission
i) To study how serum drug levels and/or anti-drug antibodies are associated to drug efficacy in RA remission
j) To assess the performance of definitions of RA flare
k) To assess differences in success of DMARD reduction and discontinuation in patients according to disease duration

## 3. Study population

### *3.1 Overview*

The main study population will consist of adult men and women with RA according to the 2010 ACR/EULAR classification criteria (51) (Appendix 2) and sustained remission for a year. In addition, patients who participated in the first ARCTIC trial will be included.

### *3.2 Number of patients and assignment to treatment groups*

One hundred and sixty patients using TNFi (Group B) and 160 patients using synthetic DMARDs (Group S) will be randomized. Approximately 40 patients with active disease from the first ARCTIC study are estimated to be included for longitudinal follow-up in ARCTIC REWIND, and these patients might switch to Group B or Group S if they become eligible for inclusion into these groups during follow-up. Altogether, 360 patients will be included in the study. For sample size calculations, see 7.3 Sample size calculations, page 48.

### *3.3 Inclusion criteria*

In order to participate in this study the subjects must meet the following inclusion criteria:

a) Rheumatoid arthritis according to the 2010 ACR/EULAR classification criteria (52), Appendix 2.
b) Male or non-pregnant, non-nursing female
c) >18 years of age and <80 years of age
d) Patients in Group B (TNFi): Any disease duration. Patients in Group S (synthetic DMARDs): RA diagnosis after 01.01.2010.
e) Sustained remission for ≥12 months according to DAS or DAS28, with documented remission status at a minimum of 2 consecutive visits during the last 18 months **OR** participation in the first ARCTIC trial
f) DAS <1.6 and no swollen joints at inclusion **OR** participation in the first ARCTIC trial
g) Unchanged treatment with TNFi and/or synthetic DMARDs during the previous 12 months, with a stable or reduced dose of glucocorticosteroids **OR** participation in the first ARCTIC trial
h) Subject capable of understanding and signing an informed consent form
i) Provision of written informed consent

### *3.4 Exclusion criteria*

Patients with any of the following criteria will not be eligible to participate in the study:

a) Abnormal renal function, defined as serum creatinine >142 µmol/L in female and >168 µmol/L in male, or a GFR <40 mL/min/1.73 $m^2$
b) Abnormal liver function (defined as ASAT/ALAT >3x upper normal limit), active or recent hepatitis, cirrhosis
c) Major co-morbidities, such as severe malignancies, severe diabetic mellitus, severe infections, uncontrollable hypertension, severe cardiovascular disease (NYHA class 3 or 4) and/or severe respiratory diseases
d) Leukopenia and/or thrombocytopenia
e) Inadequate birth control, pregnancy, and/or breastfeeding
f) Indications of active TB
g) Psychiatric or mental disorders, alcohol abuse or other substance abuse, language barriers or other factors which makes adherence to the study protocol impossible.

# 4. Investigational plan

## *4.1 Overview of study design*

ARCTIC REWIND is a 36-month, randomized, open-label, phase IV, parallel-group, multi-center, non-inferiority study to evaluate the effect of DMARD dose reduction in RA subjects who have achieved sustained remission. The design of the trial is illustrated in **Figure 1**.

All RA patients followed at the study centers who are in sustained (≥12 months) remission (defined as DAS28 <2.6 or DAS <1.6) are potential study patients. In addition, patients who were participants in the first ARCTIC trial but have active disease or less than 12 months of remission will be followed in ARCTIC REWIND. The selection criteria for ARCTIC REWIND are presented in section *3.3 Inclusion criteria* and *3.4 Exclusion criteria*.

Eligible patients with informed consent will be separated into three groups:
**Group B:** Patients in sustained remission receiving a TNFi (**B**iologic), either as monotherapy or in combination with a synthetic DMARD
**Group S:** Patients in sustained remission receiving **S**ynthetic DMARD therapy (either monotherapy or combination therapy) without concomitant biologic DMARD therapy
**ARCTIC follow-up**: Patients from the first ARCTIC study with active disease or less than 12 months remission

Patients included in ARCTIC REWIND will be followed for 36 months, with visits every 4 months. The study is divided into three periods as outlined in **Table 1.**

Each study centre will have a phone number for patients to call in case of increasing disease activity. If a patient is experiencing a potential disease flare, a visit will be arranged within one week to allow for a thorough examination and documentation of disease status. If a disease flare occurs (see section *4.3.1 Flare definition*), the patient will return to the full dose study medication (the treatment the patient was receiving at baseline).

**Figure 1:** Illustration of ARCTIC REWIND study design.

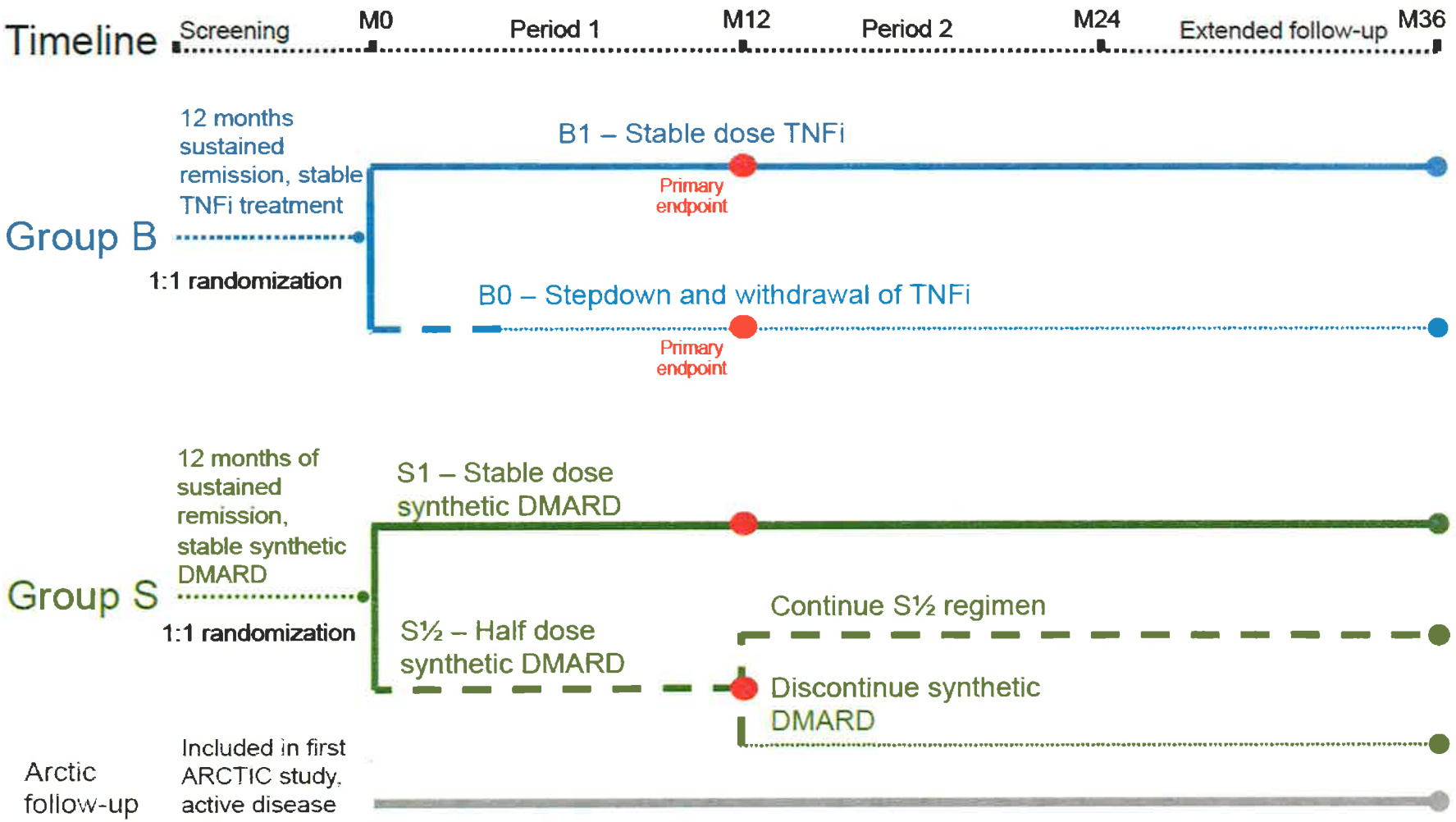


**Table 1:** Overview of ARCTIC REWIND treatment periods

| | *Period 1*<br>*Baseline (Visit 1) to 12 months (Visit 4)* | *Period 2*<br>*12 months (Visit 4) to 24 months (Visit 7)* | *Period 3*<br>*24 months (Visit 7) to 36 months (Visit 10)* |
|---|---|---|---|
| **Group B** | Patients in **Group B** will be randomized 1:1 into continued stable TNFi treatment (arm B1) or four months of half dose TNFi therapy followed by discontinuation of the TNFi (arm B0). The reduction in TNFi dose during the first four months in arm B0 will take place by either reducing the dose by 50% without changing the treatment intervals, or by doubling the treatment intervals without changing the TNFi dose, as outlined for each drug in **Table 2.** Any co-medication with synthetic DMARDs will be kept stable throughout the study. | **Group B** patients in sustained remission will continue to receive the treatment they were assigned to in Period 1. Concomitant synthetic DMARD therapy is kept stable. Patients are followed with regards to efficacy and safety measures. | **Group B** patients in sustained remission will continue to receive the treatment they were assigned to in Period 1. Concomitant synthetic DMARD therapy is kept stable. Patients are followed with regards to efficacy and safety measures. |
| **Group S** | Patients in **Group S** will be randomized 1:1 into continued stable synthetic DMARD(s) (S1) or half dose synthetic DMARD(s) (S½). The reduction in synthetic DMARD dose will be done according to the plan in **Table2.** | **Group S** patients in sustained remission on **S1** regimen (stable dosage) will continue this treatment. **Group S** patients in sustained remission on S½ treatment will be randomized into continued half dose synthetic DMARD treatment (S½→½) or discontinued synthetic DMARD treatment (S½→0). | **Group S** patients in sustained remission will continue to receive the treatment they were assigned to in Period 2. Patients are followed with regards to efficacy and safety measures. |
| **ARCTIC follow-up** | Patients in the **ARCTIC follow-up group** will not be randomized at the start of the study, but evaluated with efficacy and safety measures at each visit and continue to receive treatment according to the predefined ARCTIC protocol. If the patient becomes eligible for inclusion in Group S or Group B, the patient can be re-included and randomized in the appropriate group. | Same as for Period 1. | Same as for Period 1. |

### *4.2 Dosage and Drug Administration*

*4.2.1 Group B*

Patients in Group B will at baseline either use a TNFi (etanercept, adalimumab, infliximab, certolizumab pegol, golimumab) in monotherapy, or a TNFi in combination with a synthetic DMARDs. Any co-medication with synthetic DMARDs should be kept stable throughout the study. At baseline patients in Group B will be randomized to either continued full dose of the TNFi (B1) or step-down to withdrawal of the TNFi (B0). The B0 group will in the period from baseline to 4 months receive half dose of the TNFi, and the TNFi will then be discontinued from month 4 if no flare has occured (see also **Figure 1** and **Table 1**). The half dose regimen for each TNFi is outlined in **Table 2:** Dosing regimens in ARCTIC REWIND.

Patients randomized to B1 at baseline will continue this treatment regimen throughout the study as long as the treatment result is satisfactory. If a patient in the treatment arm B1 experiences a flare, treatment can be escalated according to the physician's preference. The ARCTIC treatment guidelines (*Appendix 14*) can be used as a tool to help make this decision.

If a patient in arm B0 experiences a flare, he or she will return to the full dose of the study medication, i.e. the treatment the patient was receiving at baseline.

*4.2.1 Group S*

Patients in Group S will at baseline either receive a synthetic DMARD in monotherapy, or more than one synthetic DMARD in combination therapy. Patients entering Group S cannot have received any biologic DMARDs during the previous 12 months. At baseline patients in Group S will be randomized to either continued stable synthetic DMARD treatment (S1) or half dose synthetic DMARD treatment (S½). Patients on combinations of synthetic DMARDs who are randomized to the S½ arm will continue combination therapy, but each DMARD will be reduced to half dose. The half dose regimen for each synthetic DMARD is outlined in **Table 2:** Dosing regimens in ARCTIC REWIND.

In Period 2, patients in arm S1 who have not had a flare will continue their treatment unchanged. Patients in arm S½ who have not experienced a flare will be randomized 1:1 to either continue the S½ treatment or discontinue their synthetic DMARD(s). These patients will at this second randomization have been in sustained remission for at least two years.

Patients randomized to S1 at baseline will continue this treatment regimen throughout the study as long as the treatment result is satisfactory. If the patient in the S1 arm experiences a flare, treatment can be escalated according to the physician's preference, preferentially as outlined in the ARCTIC treatment guidelines (Appendix 16.14).

If a patient in treatment arms S½ or S0 experiences a flare, he or she will return to the full dose of the study medication (the treatment the patient was receiving at baseline).

*4.2.1 ARCTIC follow-up group*

Patients in the ARCTIC follow-up group will continue to receive DMARD therapy targeted towards clinical remission as defined in the ARCTIC trial (DAS<1.6 and no swollen joints) and based on the ARCTIC treatment guidelines (Appendix 16.14 and 16.15).

**Table 2:** Dosing regimens in ARCTIC REWIND

| *Drug* | *Producer* | *Package* | *ATC* | *Standard full dosage* Dosage | Frequency | *1/2 dosage regimen* Dosage | Frequency |
|---|---|---|---|---|---|---|---|
| **Synthetic DMARDs** | | | | | | | |
| **Methotrexate tablets** *(methotrexate)* | Pfizer | 2,5 mg no.100 | L04A X03 | 25mg p.o. | Weekly | 12,5mg p.o. | Weekly |
| **Metex prefilled syringe** *(methotrexate)* | Medac | 50mg/ml 0,5ml x 6 syringes | L01B A01 | 25mg s.c. | Weekly | | |
| **Metex prefilled syringe** *(methotrexate)* | Medac | 50mg/ml 0,25ml x 6 syringes | L01B A01 | | | 12,5mg s.c. | Weekly |
| **Salazopyrin EN tablets** *(sulfasalazine)* | Pfizer | 500mg No. 100 | A07E C01 | 1000mg x 2 p.o. | Daily | 500mg x 2 p.o. | Daily |
| **Plaquenil** *(hydroxychlorochine)* | Sanofi-Aventis | 200mg No. 100 | P01B A02 | 200mg x 2 p.o. | Daily | 200mg x 1 p.o. | Daily |
| **Arava** *(leflunomide)* | Sanofi-Aventis | 20mg No 100 | L04A A13 | 20mg p.o. | Daily | | |
| **Arava** *(leflunomide)* | Sanofi-Aventis | 10mg No 100 | L04A A13 | | | 10mg p.o. | Daily |
| **TNF inhibitors** | | | | | | | |
| **Enbrel** *(etanercept)* | Pfizer | 25 mg No. 4 sets of solution + syringe | L04A B01 | 50 mg s.c. | Weekly | 25 mg s.c. | Weekly |
| **Cimzia** *(Certolizumab pegol)* | UCB | 200 mg 2 x 1ml pre-filled syringe | L04A B05 | 200mg s.c. | Bi-weekly | 200mg s.c. | Every 4 weeks |
| **Simponi** *(Golimumab)* | Janssen Biologics | 50 mg 1 prefilled syringe | L04A B06 | 50 mg s.c. | Every 4 weeks | 50 mg s.c. | Every 8 weeks |
| **Remicade** *(infliximab)* | Janssen Biologics | 100 mg No. 1 solution | L04A B02 | 3-5 mg/kg i.v. | Every 8 weeks | 1.5-3 mg/kg i.v. | Every 8 weeks |
| **Humira** *(adalimumab)* | AbbVie | 40 mg/0.8 ml No. 2 x 0.8ml prefilled syringe | L04A B04 | 40 mg s.c. | Bi-weekly | 40 mg s.c. | Every four weeks |

### *4.3 Efficacy and safety assessments*

Efficacy and safety will be evaluated at each visit with the variables described in section 5.3, 5.4, 5.5, 5.6, 5.7 and in Appendix 16.1

#### *4.3.1 Definition of flare*

A flare is defined as an increase in DAS to >1.6 with a change in DAS of ≥0.6 and more than one swollen joint based on examination of 44 joints (see *Appendix 4* for description of joint examination). If a patient does not fulfill this formal definition, but experiences a clinically significant flare according to the investigator and patient, this should be treated as a flare but recorded separately in the CRF.

### *4.4 End of study*

The study period is 36 months, with a possible extension to 60 months which will be decided by the study committee. The end of the trial is defined as the date of the last visit of last participating patient in this study.

The possible extension to 60 months will further assess both primary and secondary outcomes, with a focus on structural damage, physical function and occurrence of extra-articular manifestations, as well as economic evaluations, and will necessitate an amendment of the protocol which will be assessed by the regional ethics committee.

### *4.5 Concomitant medication (except intra-articular injections)*

Concomitant medication will be recorded in the CRF, with particular attention to registration of concomitant antirheumatic and pain medications.

NSAIDs including COX-2 selective inhibitors (coxibs) are permitted. The choice and dosage of NSAIDs/coxibs will be at the discretion of the treating rheumatologist and shall be recorded in the CRF. Analgesics up to the maximum recommended dose may be used for pain relief as required. Patients should avoid analgesics within 24 hours prior to a visit if possible.

All patients will receive vitamin D and calcium supplements as appropriate according to glucocorticoid treatment, and postmenopausal women and older men (>70 year) will be considered for bisphosphonate treatment according to general guidelines. Oral glucocorticoids

at inclusion are allowed in doses of maximum 5 mg, and should be stable or reduced during the last year. Intramuscular or intravenous glucocorticoids are not allowed as long as the patient is receiving the study medication. If a patient experiences a flare, the flare might be treated with other DMARDs than the study drug(s) or intramuscular, intravenous or oral glucocorticoids.

### *4.6 Intra-articular injections*

Patients can receive intra-articular injections in one swollen joint at each visit and still be classified as not having a flare. Any injections can be ultrasonography guided if deemed necessary by the physician. Patients who experience a flare can receive intra-articular injections as needed, up to a maximum dosage of 80mg triamcinolone hexacetonide per visit.

### *4.7 Subject compliance*

Compliance to treatment will be assessed retrospectively at each visit by asking the subject to report their compliance as either full, partial or low. Full compliance is defined as taking the treatment as prescribed, partial compliance is less than full but more than 70% compliant to prescribed treatment, and low compliance is less than 70% compliant.

### *4.8 Drug storage*

Each investigator is responsible for ensuring that the patient knows the correct storage of the study drug(s) and adheres to this.

### *4.9 Drug labelling*

Drub labeling will be performed according to "GUIDE TO GOOD MANUFACTURING PRACTICE FOR MEDICINAL PRODUCTS ANNEXES", annex 13, 1st September 2009 and the Norwegian "Veiledning til Forskrift av 30. oktober 2009 om klinisk utprøving av legemidler til mennesker".

### *4.10 Schedule modifications*

Protocol modifications to an ongoing study must be approved by the study committee, and must be approved by the National Medical Leader and biostatistician. Protocol modifications will be submitted to the REK for approval and to the regulatory agencies (SLV) as required.

### *4.11 Dose modifications, interruptions and delays*

The choice of drug and dosing regimens should be according to Table 2. Other approved dosing regimens and drugs with same ATC code and Norwegian marketing authorization may be acceptable if deemed important to the patient by the treating physician. Guidelines for modification of treatment regimens based on abnormal laboratory values and/or adverse events are outlined in **Table 3**. If clinical situations occur that are not described in **Table 3**, the physician should make treatment decision based on the summary of product characteristics (SPC), clinical judgment and if necessary contact with the Central Study Coordinating Team (CSCT).

**Table 3:** Abnormal laboratory values and/or adverse events - dose modifications

| *Condition* | *Action* |
|---|---|
| **1. Anemia** | Policy after analyses according to the rheumatologists own insight |
| **2. Leucopenia** | • Number of white blood cells 2.5 - 3.5 *$10^9$ / L: **Stop MTX until recovery. Thereafter, half dose MTX**<br>• Number of white blood cells ≤2.5 *$10^9$ / L: **Stop MTX** |
| **3. Thrombocytopenia** | • Platelet count 100 - 150 *$10^9$ / L: **Stop MTX until recovery. Thereafter, half dose MTX**<br>• Platelet count ≤100 *$10^9$ / L: **Stop MTX** |
| **4. Pancytopenia** | • White blood cell (WBC) count ≤2.5 * $10^9$ / L and platelet count ≤100 * $10^9$ / L: **Stop MTX** |
| **5. Severe nausea and/or dyspepsia** | **Change to MTX subcutaneously** |
| **6. Oral ulcers** | Very severe: **Stop MTX until recovery, thereafter, half dose MTX** |
| **7.Transaminase increase** | • ≧2x upper normal limit: **Do not increase MTX**<br>• ≧3x upper normal limit: **Stop MTX until recovery, thereafter, half dose MTX**<br>• ≧3x upper normal limit and additional signs of liver toxicity (bilirubin >2x upper normal limit, INR >1.5x upper normal limit, alkaline phosphatase >2x upper normal limit, presence of worsening fatigue, nausea, vomiting, fever, rash or eosinophilia): **Stop MTX** |
| **8. Renal function** | GFR decreased by 25%: **Stop MTX until recovery, thereafter administration of half dose MTX** |
| **9. Pneunonitis** | In case of pneumonitis caused by MTX: **Stop MTX** |
| **10. Other** | Possible other side effects for synthetic DMARDs and TNFi will be at the treating rheumatologist's discretion, according to established guidelines. |

### *4.12 Procedures for discontinuation*

4.12.1 Patient discontinuation

Patients have the right to withdraw from the study at any time for any reason. In the case that a patient decides to prematurely withdraw from the study, he or she should be asked if they can still be contacted for further information, so that a final evaluation can be made with an explanation of why the patient is withdrawing from the study, including assessment of possible adverse events. Although a subject is not obliged to give his or 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.

If possible, at the last visit of the patient, all assessments of the "End of study visit" will be done.

The investigator is obliged to follow up any significant adverse events until the outcome is either recovered or resolved, recovering or resolving, not recovered or not resolved, recovered or resolved with sequelae, fatal or unknown.

4.12.2 Treatment discontinuation

Patients who for some reason have to withdraw from the study medication will be asked to continue follow-up in the study

4.12.3 Trial discontinuation

The study committee reserves the right to terminate the study at any time. This may be due to safety reasons or if new knowledge arises that invalidates the study (including results from interim analyses). Other reasons that may have a major impact on the study, including ethical and financial aspects, and difficulties in the recruitment of patients, may also lead to termination of the study. In terminating the study, the study committee and investigators will assure that adequate consideration is given to the protections of patients' interests. The sponsor and principal investigator will inform all investigators and the relevant regulatory authorities of the termination of the trial along with the reasons for such action. If the study is terminated early on grounds of safety, the relevant authorities should be informed within 15 days.

### *4.13 Rationale for study design*

Remission is the main goal of RA treatment, and studies of modern treatment regimens report a majority of patients reaching remission. As these treatment strategies are implemented in clinical practice, physicians will experience increasing numbers of RA patients in sustained remission. There are no current recommendations on how to best treat these patients, and whether DMARDs can be tapered or withdrawn.

The study is designed to test whether DMARD dose reduction is an alternative in sustained RA remission, and test DMARD withdrawal strategies in a population of patients who have been in sustained remission for a prolonged time and show no clinical signs of inflammation at baseline.

#### 4.13.1 Rationale for treatment alternatives

Previous studies imply that DMARD dose reduction is possible in RA low disease activity, and that some patients can discontinue DMARD therapy in RA remission. The treatment alternatives outlined in the current design for ARCTIC REWIND were chosen based on review of literature, discussion with international experts and discussion with Norwegian clinicians and researchers.

### *4.14 Linkage to other registers*

In addition to the variables outlined in Table 2, patients will be asked to give consent to collection of data from registries such as The Norwegian Prescription Database, The Norwegian Health Economics Administration, and the Norwegian Patient Register. This will allow certain outcomes to potentially be obtained though linkage to national medical or public registers and databases to answer research questions related to safety and health economics. Examples of such outcomes are cancer and other serious adverse events, health care utilization, work participation and social benefits. The Cancer Registry of Norway (Kreftregisteret), the Norwegian Patient Registry (Norsk pasientregister – NPR), the Cause of Death Registry (Dødsårsaksregisteret), the Norwegian Prescription Database (Reseptregisteret), the Norwegian Myocardial Infarction Register (Norsk hjerteinfarktregister), the Norwegian Surveillance System for Communicable Diseases (Meldingssystem for smittsomme sykdommer – MSIS) and The Norwegian Labour and Welfare Administration (NAV) are potential data sources. The patient consent form includes information about linkage.

Participation in international collaboration involving sharing of data from the ARCTIC REWIND study and merging of ARCTIC REWIND data with other (similar) studies will be based on fully de-identified data.

### *4.15 Centers*

13 centers have agreed to participate in this study, aiming to recruit 10 - 100 patients per site. These centers are:

Diakonhjemmet Hospital (coordinating center and sponsor)
Universitetssykehuset i Nord-Norge
Helse Møre og Romsdal HF, Ålesund sjukehus
Haukeland Universitetssykehus Bergen
Sørlandet Sykehus Kristiansand HF
Haugesund Sanitetsforenings Revmatismesykehus AS
Martina Hansens Hospital
Sykehuset Østfold HF Moss
Lillehammer Revmatismesykehus
Vestre Viken HF Drammen Sykehus
Privatpraksis Bendvold/Dovland Kristiansand
Revma Vestfold
Helgelandssykehuset Mo i Rana

More centers might be invited to participate as long as inclusion is still open.

## 5. Study procedures

The visits will be carried out according to the schedule shown in *Appendix 1* and data collection includes swollen joint assessments of 44 joints and Ritchie articular index, laboratory tests, vital signs, standardized assessments of health related quality of life and physical function and imaging procedures. Standardized ultrasonography examinations will be performed yearly in all patients. Conventional radiographs (CR) of hands and feet and MRI of metacarpophalangeal joints and wrist of dominant hand will be acquired at baseline, 12, 24 and 36 months. Scoring of the images (except for ultrasonography) will be performed by trained experts at Diakonhjemmet Hospital or cooperating institutions based on methods which have met internationally validated standards.

The sequence of assessments and procedures at each visit (as specified in Appendix 1) is to be standardized as follows:

1. Laboratory samples must be drawn at least 1 hour prior to physician assessments, and up to 7 days in advance
2. Patient reported health outcomes assessments: VAS for pain, global assessment of diseases status and fatigue, HAQ-PROMIS, SF-36, EQ-5D, RAID, WPAI, OMERACT Preliminary Flare Questionnaire.
3. Study nurse/investigator assessments:
    - Joint counts, preferably done by the same assessor for consistency
    - Registration of co-medication
    - Safety assessments (AEs, vital signs)
    - Nurse/investigator global assessment of disease activity (VAS)
4. CR and MRI when applicable (this can alternatively be done after 5.)
5. Treating physician:
    - Review of laboratory data
    - US at yearly visits

### *5.1 Screening examination and eligibility screening form*

All patients must sign and date the written informed consent before any study specific assessments or procedures are performed. A screening examination should be performed

before the start of the study (or at the baseline visit for patients that fulfill the entry criteria). The following procedures and assessments have to be completed:

- Physical examination including pulse rate and blood pressure
- Laboratory tests (CRP, ESR, hematology, blood chemistry, urine dipstick)

### *5.2 Procedures for enrolment of eligible patients*

Eligible patients will be randomized through a pre-defined procedure, stratified for site and previous participation in ARCTIC (if applicable). Randomization will be performed electronically through the eCRF (Viedoc).

### *5.3 End of study visit*

The end of study visit should contain the examinations planned for visit 10 (month 36), including radiographs, MRI and ultrasonography if at all achievable, see Appendix 1 for details.

### *5.4 Extra visit in case of flare*

If the patient suspects a flare in disease activity, he or she should contact the study site immediately and be seen there within a week. Examinations as outlined in Appendix 1 should be performed, including ultrasonography. If a flare of disease activity according to the definition in section *4.3.1 Definition of flare* is confirmed, treatment should be modified as outlined in sections *4.1* and *4.2*.

### *5.5 Clinical assessments and procedures*

#### 5.5.1 Efficacy

Efficacy variables include individual parameters of the ACR core set of disease activity measures, ACR responses, DAS, EULAR response, the van der Heijde modified Sharp score of radiographs, the 32 joint PD and grey scale US score and RAMRIS (plus Haavardsholm tenosynovitis score and the OMERACT MRI joint space narrowing score). Details regarding efficacy variables are outlined below.

ACR core set

Tender and swollen joint counts, VAS scales for pain, patient and investigator global assessment of disease activity, patient-assessed disability (HAQ) and acute phase response (ESR/hsCRP).

ACR response

The ACR response rates ACR20, ACR50, ACR70 and ACR90 will be calculated. An ACR20/50/70/90 response is defined as ≥ 20/50/70/90 % improvement in swollen and tender joint counts plus ≥ 20/50/70/90 % improvement in 3 of the 5 remaining ACR core set components.

Disease Activity Score (DAS)

The DAS composite score includes the Ritchie articular index, the 44 swollen joint counts, the Erythrocyte Sedimentation Rate and a general health assessment on a VAS. The DAS is calculated as follows:

**DAS = 0.54*sqrt(RAI) + 0.065*(swollen44) + 0.33*Ln(ESR) + 0.0072*GH VAS**

High disease activity is defined as a DAS value >3.7, moderate disease activity as DAS >2.4 – 3.7, low disease activity as a DAS-value of 1.6 – 2.4, and remission as DAS <1.6

EULAR response

Based on the DAS, response criteria have been developed: the EULAR response criteria (Table 4). The EULAR response criteria include not only change in disease activity but also current disease activity. Three categories are defined: good, moderate, and non-responders (Table 4).

**Table 4:** EULAR response

| DAS at endpoint | Improvement in DAS | | |
|---|---|---|---|
| | >1.2 | >0.6 and ≤1.2 | ≤0.6 |
| ≤2.4 | Good | Moderate | None |
| >2.4 and ≤3.7 | | | |
| >3.7 | | | |

FDA Major clinical response

Requires achieving ACR70 response for at least 6 subsequent months (ACR70 response means 70% improvement of tender and swollen joint count coupled with 70% improvement in

3 of 5 of the following: patient's assessment, physician's assessment, ESR or CRP, pain scale, Health Assessment Questionnaire).

**Remission definitions**

ACR remission criteria, ACR and EULAR remission criteria, and remission criteria based on DAS and DAS28 as well as the US Food and Drug Administration (FDA) definition of remission will be calculated (53).

ACR and EULAR 2011 remission criteria

The ACR and EULAR 2011 Boolean remission criteria state that all of the following must be present (13;14) :

- CRP ≤ 1 mg/dL
- SJC ≤ 1
- TJC ≤ 1
- Patient global assessment of disease activity VAS ≤ 1 on a scale from 0-10

1981 ACR remission criteria

According to the ACR remission criteria, a minimum of five of the following items must be present for at least two subsequent months:

- Morning stiffness <15 minutes
- No fatigue
- No joint pain by history
- No joint tenderness or pain on motion
- No soft-tissue swelling in joints or tender sheaths
- ESR <30 mm/1st hour in women or <20 mm/1st hour in men

Disease activity score remission (DAS) criteria

DAS remission defined as a score <1.6 using a composite index of the following measures:

- Ritchie articular index of tender joints
- 44 swollen joint count
- ESR
- Patient's assessment of general health (measured on a 100 mm visual analogue scale)

DAS28 remission criteria

DAS28 remission defined as a score <2.6 using a composite index of the following measures:

- 28-joint count for tender and swollen joints
- ESR
- Patient's assessment of general health

FDA Remission criteria

Requires achieving ACR clinical remission and absence of radiological progression (Larsen or Sharp-van der Heijde method) over a continuous 6 month period in the absence of DMARDs

FDA Complete clinical remission

Same as FDA remission, but while continuing DMARD therapy

Other remission criteria

In addition, a "complete clinical" DAS remission will be calculated (similar to the FDA complete clinical remission, but for DAS) and is defined as follows:
Same as "DAS-remission" (i.e. DAS<1.6), but sustained for >6 months, but requires at the same time absence of clinically swollen joints and no radiographic progression over a >6 months period.

**Imaging**

Sharp van der Heijde Score

Radiographs of hands (posterior/anterior) and foot (anterior/posterior) will be taken at baseline, 12, 24 and 36 months. The modified Sharp van der Heijde Score (vdHSS) will be calculated, including an erosion score and a joint space narrowing score.

According to the vdHSS, erosion is assessed in 16 joints (five MCP, four PIP, IP of the thumbs, 1st MCB, radius and ulna bones, trapezium and trapezoid as one unit (multangular), navicular, lunate) for each hand and wrist, and six joints (five MTP, IP) for each foot. One point is scored if erosions are discrete, rising to 2, 3, 4, or 5 depending on the amount of surface area affected. The score for erosion ranges from 0 to 160 in the hands and from 0 to 120 in the feet (the maximum erosion score for a joint in the foot is 10). JSN is assessed in 15 joints (five MCP, four PIP, CMC 3 to 5, multangular navicular-lunate, radiocarpal) for each hand and wrist, and six joints (five MTP, IP) for each foot. JSN is combined with a score for (sub)luxation

and scored as follows: 0 = normal; 1 = focal or doubtful; 2 = generalized, less than 50% of the original joint space; 3 = generalized, more than 50% of the original joint space or subluxation; 4 = bony ankylosis or complete luxation. The score for JSN ranges from 0 to 120 in the hands and from 0 to 48 in the feet (54). The total score is the sum of scores of erosion and JSN, the maximum score being 448.

Radiographs will be scored at a central reading site by two experienced readers blinded to patient identity, clinical information and timepoints of radiographs, grouped per patient.

Ultrasonography score

A Siemens Antares (or machine with similar specifications) will be used for the ultrasonography assessment, with the following hardware specifications:
*PRF = 391, Frequency = 7.3 MHz, R/S = 5, Filter = 2*

The joints outlined below will be scored for grey scale synovitis and power Doppler signal, with the best possible match to a reference atlas developed by dr. Hilde Berner Hammer (55). The scoring system is based on the OMERACT US recommendations, with semi-quantitative scores from 0 - 3 in 32 joints: Bilateral MCP I-V, RCJ, DRUJ, intercarpal, elbow, knee, talo-crural joint and MTP I-V, yielding a maximum total score of 64. In addition to the 32 joints the following will be scored: Bilateral PIP2 and 3 joints, extensor carpi ulnaris tendon and tibialis posterior tendon.

US examination will be performed in all patients at 0, 12, 24 and 36 months, and at visits where a flare is suspected or detected.

MRI (RAMRIS, tenosynovitis score, joint space narrowing)

Magnetic Resonance Imaging (MRI) of the dominant hand and wrist pre- and post-gadolinium will be performed at baseline, 12 and 24 months. MRI at 36 months will be performed without gadolinium-based contrast agent. The same side will be assessed at all time-points.

MR Images will be read according to the RAMRIS, tenosynovitis and joint space narrowing score at a central reading site by two experienced readers. The RAMRIS consists of MRI definitions of important joint pathologies, a core set of MRI sequences and a semi-quantitative scoring system for erosions, bone marrow edema and synovitis (56). The RAMRIS core set of

MRI sequences to assess inflammatory as well as destructive changes in RA joints includes: Imaging in 2 planes, with T1-weighted images before and after intravenous gadolinium-contrast to assess synovitis and erosions, plus a T2-weighted fat saturated sequence or a STIR sequence to assess bone marrow edema.

Both flexor and extensor tenosynovitis will also be evaluated, according to the scoring system described by Haavardsholm et al, at the level between the radioulnar joint and the hook of the hamate, thus including both wrist and finger tendons (57). Flexor and extensor tenosynovitis are evaluated semi-quantitatively in 10 different anatomical areas, graded from grade 0 to 3 (total score 0–30). The grading is based on the maximum width (in mm) of post-contrast enhancement within each anatomical area on axial T1-weighted MR images.

Joint space narrowing will be read according to the latest scoring system released by the OMERACT group, a provisional definition grades MCP joints and wrist joints 0-4 according to loss of inter-bone distance (58).

5.5.2 Safety

Physical examination

A general physical examination (including the cardiovascular, respiratory, GI and neurological systems) should be performed at each visit, and recorded as normal or abnormal, with a description of abnormalities. Diagnosis of new abnormalities, or worsening of abnormalities, should be recorded as an AE if appropriate.

Vital signs

Vital signs including pulse rate, systolic and diastolic blood pressure and body weight and height will be assessed at times indicated in table 3.

### *5.6 Laboratory assessments*

The following laboratory tests will be recorded at the time points indicated in Appendix 1

Hematology / complete blood count

Hemoglobin, hct, erythrocytes, white blood cells with differentials, platelet counts.

Blood chemistry

AST and/or ALT, ALP, albumin, creatinine, random glucose, potassium, sodium,

Urinalyses

Dipstick for blood, protein and glucose (with microscopic examination as clinically indicated).

Acute phase reactants

CRP is measured by high sensitivity CRP nephelometry (mg/dL) and ESR by the Westergren method (mm/hr), according to local practice.

### *5.7 Quality of life, disability and utility assessments*

SF-36

The SF-36 is a multi-purpose, short-form health survey with 36 questions (59). It yields an 8-scale profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures and a preference-based health utility index (SF-6D) (60). It is a generic measure, as opposed to one that targets a specific age, disease, or treatment group. Accordingly, the SF-36 has proven useful in surveys of general and specific populations, comparing the relative burden of diseases, and in differentiating the health benefits produced by a wide range of different treatments.

Health Assessment Questionnaire

The Stanford Health Assessment Questionnaire (HAQ) was introduced in the 1980s and is now widely used in evaluation of physical function in patients with RA (61). This instrument includes questions concerning the ability of patients to perform 20 activities of daily living yielding a 0-3 score, and the measure is commonly referred to as the HAQ disability index (HAQ-DI) (62). A new version has recently been developed, the Patient-Reported Outcomes Measurement Information (PROMIS) HAQ, including a 20-item short form that will be used in this study. While the original HAQ had 4 response categories, this new version includes a fifth response option, "with a little bit of difficulty". The scoring algorithm was also changed from the HAQ's 0–3 unit scale to a 0–100 unit scale. Completion time was reduced by over one-third compared to the original HAQ.

EQ-5D

EQ-5D is a utility instrument for measurement of health related quality of life (63;64) Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and a single index value for health status.

RAID

The RA Impact of Disease (RAID) score is a patient-derived composite response index for use in clinical trials in RA. It includes seven domains with the following relative weights: pain (21%), functional disability (16%), fatigue (15%), emotional well-being (12%), sleep (12%), coping (12%) and physical well-being (12%) (65).

OMERACT Preliminary Flare Questionnaire

The OMERACT RA Flare Working Group has developed a provisional questionnaire (Appendix 16) including information about variables deemed to be of relevance by both patients and health care professionals to describe and define RA flare. The questions include assessments of severity and duration of flare, self-management strategies related to flare, pain, function, fatigue, stiffness, participation, coping, and patients' self-assessment of joint tenderness and swelling.

### *5.8 Worker productivity*

Worker productivity is generally subdivided into 2 components: absenteeism and presenteeism. The concept of absenteeism has been defined as productivity loss due to health-related absence from work, while presenteeism refers to reduced performance or productivity while at work due to health reasons. Absenteeism may include personal time off, sick days off work, time on short and/or long-term work disability, or time on worker's-compensated days; and presenteeism could be characterized as the time not being on the task, or decreased work quality and quantity. Patients will be asked to answer the Work Productivity and Activity Impairment Questionnaire: Rheumatoid arthritis V2.0 (WPAI:RA).

The WPAI yields four types of scores:

1. Absenteeism (work time missed)
2. Presenteeism (impairment at work/reduced on-the-job effectiveness)
3. Work productivity loss (overall work impairment / absenteeism plus presenteeism)
4. Activity Impairment

### *5.9 Genetic markers and biomarkers*

Samples (including serum, plasma, full blood and urine) for biomarker or DNA/RNA discovery and validation will be collected and stored in a freezer at -70° C at visits shown in table 3. These samples will be used for research purposes only, and may include measurement of cytokines and other known or potential new markers of inflammation or damage, such as interleukins, interferons, metalloproteases, transforming growth factor, TNFs, adhesion molecules etc as well as DNA/RNA analyses (genomics and proteomics). All samples will be stored in a certified biobank. Some analyses might take place in other countries if necessary.

## 6. Adverse events

The investigator is responsible for the detection and documentation of events meeting the criteria and definition of an adverse event (AE) or serious adverse event (SAE). Each patient will be instructed to contact the investigator immediately should they manifest any signs or symptoms they perceive as serious. The methods for collection of safety data are described below.

### *6.1 Definitions*

6.1.1 Adverse events

An AE is any untoward medical occurrence in a patient administered a pharmaceutical product and which does not necessarily have a causal relationship with this treatment. An adverse event (AE) can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product.
The term AE is used to include both serious and non-serious AEs.

6.1.1 Serious Adverse Event (SAE)

Any untoward medical occurrence that at any dose:

- Results in death
- Is immediately life-threatening
- Requires in-patient hospitalization or prolongation of existing hospitalization
- Results in persistent or significant disability or incapacity

- Is a congenital abnormality or birth defect
- Is an important medical event that may jeopardize the subject or may require medical intervention to prevent one of the outcomes listed above

6.1.2 Suspected Unexpected Serious Adverse Reaction (SUSAR)

Adverse Reaction: all untoward and unintended responses to an investigational medicinal product related to any dose administered.

Unexpected Adverse Reaction: an adverse reaction, the nature or severity of which is not consistent with the applicable product information.

Suspected Unexpected Serious Adverse Reaction: Unexpected Adverse Reaction that:

- Results in death
- Is immediately life-threatening
- Requires in-patient hospitalization or prolongation of existing hospitalization
- Results in persistent or significant disability or incapacity
- Is a congenital abnormality or birth defect
- Is an important medical event that may jeopardize the subject or may require medical intervention to prevent one of the outcomes listed above.

### *6.2 Time Period and Frequency of Detecting AE and SAE*

The standard time period for collection and recording of AE and SAEs will begin at baseline for each patient. During the course of the study all AEs and SAEs will be proactively followed up for each patient; events should be followed up to resolution, unless the event is considered by the investigator to be unlikely to resolve due to the underlying disease. Every effort should be made to obtain a resolution for all events, even if the events continue after discontinuation/study completion.

### *6.3 Recording of Adverse Events*

If the patient has experienced adverse event(s), the investigator will record the following information in the CRF:

- The nature of the event(s) will be described by the investigator in precise standard medical terminology (i.e. not necessarily the exact words used by the patient)

- The duration of the event will be described in terms of event onset date and event ended data
- The intensity of the adverse event will be described according to Common Terminology Criteria for Adverse Events version 4.0 (CTCAE)
- The Causal relationship of the event to the study medication will be assessed as one of the following

  **Unrelated**

  There is not a temporal relationship to investigational product administration (too early, or late, or investigational product not taken), or there is a reasonable causal relationship between non-investigational product, concurrent disease, or circumstance and the AE.

  **Unlikely**

  There is a temporal relationship to investigational product administration, but there is not a reasonable causal relationship between the investigational product and the AE.

  **Possible**

  There is reasonable causal relationship between the investigational product and the AE. Dechallenge information is lacking or unclear.

  **Probable**

  There is a reasonable causal relationship between the investigational product and the AE. The event responds to dechallenge. Rechallenge is not required.

  **Definite**

  There is a reasonable causal relationship between the investigational product and the AE.

- Action taken
- The outcome of the adverse event – whether the event is resolved or still ongoing.

### *6.4 Reporting Procedure*

#### 6.4.1 AEs and SAEs

All adverse events and serious adverse events that should be reported as defined in section 8.1.1 will be recorded in the patient's CRF.

SAEs must be reported by the investigator to the sponsor as outlined in the ISF, within 24 hours after the site has gained knowledge of the SAE. Every SAE must be documented by the investigator in the eCRF..The initial report shall promptly be followed by detailed, written

reports if necessary. The initial and follow-up reports shall identify the trial subjects by unique code numbers assigned to the latter.

The sponsor keeps detailed records of all SAEs reported by the investigators and performs an evaluation with respect to seriousness, causality and expectedness.

6.4.2 SUSARs

SUSARs will be reported to the Norwegian Medicines Agency and EMEAs EudraVigilance Clinical Trial Module by the sponsor according to local regulation (Veiledning til Forskrift av 30. oktober 2009 om klinisk utprøving av legemidler til mennesker). The following procedure should be followed:
The sponsor will ensure that all relevant information about SUSARs that are lethal or life-threatening is recorded and reported as soon as possible to the Norwegian Medicines Agency and EMEAs EudraVigilance Clinical Trial Module, and in any case no later than seven days after knowledge by the sponsor of such a case, and that relevant follow-up information is subsequently communicated within an additional eight days. All other suspected serious unexpected adverse reactions will be reported to the Competent Authority concerned as soon as possible but within a maximum of fifteen days of first knowledge by the sponsor.

SUSARs will be reported using the CIOMS form since Diakonhjemmet Hospital is not connected to EudraVigilance. Alternatively, the SUSAR reporting may be performed through the pharmaceutical company marketing the drug concerned.

### *6.5 Annual Safety Report*

Once a year throughout the clinical trial, the sponsor will provide the Competent Authority with a listing and a table of all suspected serious adverse reactions which have occurred over this period and a report of the subjects' safety.

### *6.6 Clinical Study Report*

The adverse events and serious adverse events occurring during the study will be discussed in the safety evaluation part of the Clinical Study Report.

### *6.7 Procedures in Case of Emergency*

The investigator is responsible for assuring that there are procedures and expertise available to cope with emergencies during the study. All patients will at inclusion receive a card with contact information about the sponsor and local investigator, as well as emergency contact

information. As described in section 5.4 all patients will receive an extra visit within one week in case of suspected flare in disease activity.

### *6.8 Pregnancy*

A female patient must be instructed to immediately inform the treating rheumatologist if she becomes pregnant during the study, which should then counsel the patient according to current guidelines. The treating rheumatologist is responsible for adjustment of treatment and any referral to other specialists necessary in case of pregnancy. The CSCT should be contacted within one working day after the investigator learns that a female study subject has become pregnant during the study.

## 7. Statistical considerations and analytical plan

### *7.1 Populations*

7.1.1 Primary (per protocol) population

Because this is a non-inferiority study, the primary population will consist of randomized patients who follow the study protocol with no major protocol violations.

Major protocol violations are defined as

- More than one missing study visit during the first 12 months of the study
- More than one period of reported low compliance during the first 12 months of the study
- Withdrawals from treatment

This population is denoted the per-protocol (PP) population. Subpopulations according to baseline DMARD treatment (Group B or S) will be used in the analyses.

Follow-up patients from the ARCTIC study without sustained remission are not randomized, and are therefore excluded from the PP population.

7.1.2 Secondary (intention to treat) population

The secondary population will consist of all randomized patients with at least one study-specific assessment after randomization. This population is denoted as the intention to treat (ITT) population. Subpopulations according to baseline DMARD treatment (Group B or S) will be used in the analyses.

Follow-up patients from the ARCTIC study without sustained remission are not randomized, and are therefore excluded from the ITT population.

7.1.3 Tertiary population

The tertiary population will consist of patients from the ARCTIC study who have not been in sustained remission for a year.

### *7.2 Primary and secondary study endpoints*

Endpoints will be defined according to study periods:

Period 1 is defined as the period from baseline to 12 months, i.e. visits 1 to 4.

Period 2 is defined as the period from 12 to 24 months, i.e. visits 4 to 7.
Period 3, the follow-up (FU) period, is defined as the period from 24 to 36 months, i.e. visits 7 to 10.
The total study period is defined as the period from baseline to 36 months, i.e. visit 1 to 10.

### 7.2.1 Primary endpoint

The primary endpoint is the proportion of patients who are non-failures at 12 months, i.e. have not experienced a flare (as defined in the study protocol, section 4.3.1) during this time period.

### 7.2.2 Secondary endpoints/exploratory endpoints include

- Proportion of patients in DAS (i.e. DAS <1.6), DAS28 (i.e. DAS28 <2.6), CDAI (i.e. ≤2.8), SDAI (i.e. SDAI ≤3.3), 2011 ACR/EULAR remission and other relevant remission criteria during the first, second, third and total study period.
- Time from randomization to first flare during the total study period
- Time in remission according to the DAS, DAS28, CDAI, SDAI and ACR/EULAR remission criteria during the total study period.
- Disease activity (status and change scores) according to DAS28, DAS, CDAI, SDAI at relevant time points during the first, second, third and total study period.
- MRI scores (RAMRIS, joint space narrowing, tenosynovitis – status and change scores) during the first, second, third and total study period.
- Ultrasound scores (grey scale synovitis, power doppler signal – status and change scores) during the first, second, third and total study period.
- Change in radiographic score (modified Sharp van der Heijde (vdHSS) during the first, second, third and total study period.
- Proportion of patients with progression of radiologic joint damage (e.g. annual increase in van der Heijde modified Sharp score ≥ 1 unit, annual increase in RAMRIS erosion score ≥ 1 unit) during the first, second, third and total study period.
- SF-36 (physical and mental component summary scores), HAQ, RAID, SF-6D and EQ-5D scores (change and status) during the first, second, third and total study period.
- Direct and indirect costs related to RA and the treatment of RA, measured during the first, second, third and total study period.

Exploratory/secondary endpoints will not be limited to those mentioned above, and will include endpoints as necessary to explore the secondary objectives of the study as described in section 2.2.

7.2.3 Safety

Safety of the treatment will be evaluated based on reported adverse events and severe adverse events, laboratory test, vital signs and performance status.

### *7.3 Sample size determination*

To estimate the expected percentage of patients who will be non-failures after one year of follow-up in ARCTIC REWIND, we used data from the Norwegian NOR-DMARD register (66). Patients who had one year of sustained DAS28 remission and were receiving TNF inhibitors in combination with MTX were selected. Of these, 80% were still in remission one year later. The corresponding number for patients receiving MTX monotherapy was 74%. Based on this, we hypothesize that also 80% of the patients included in ARCTIC REWIND who receive TNFi treatment will be in remission (an approximation of non-failure) after one year. An inferiority margin of 20% was chosen based on the belief that both clinicians and patients would accept this when keeping the high costs and potential adverse events of biologic DMARDs in mind.

If there is truly no difference between half dose and continued full dose biologic DMARD treatment on the proportion of patients in remission after 12 months, 126 patients (63 in each arm) are required in Group B to be 80% confident that the upper limit of a one-sided 97.5% confidence interval (equivalent to a 95% two-sided confidence interval) will exclude a more than 20% difference in favor of the full-dose treatment. Different combinations of remission rates and non-inferiority margins are summarized in the **Table 4**.

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 160 patients will be randomized in each of Group B and Group S. An estimated 40 patients with active disease from the ARCTIC study will be followed longitudinally without randomization in ARCTIC REWIND. The total number of patients in the study will thus be approximately 360 with an estimated recruitment period of two years.

**Table 4**: The numbers in the cells represent the total number of patients needed in each of group B and S. All calculations are based on a power of 80% and alpha 2.5%.

| Non-inferiority Margin | 90% remission at 12 m | 80% remission at 12 m | 70% remission at 12m |
|---|---|---|---|
| **15 %** | 126 | 224 | 294 |
| **20 %** | 72 | 126 | 166 |
| **25 %** | 46 | 82 | 106 |

### *7.4 Statistical and analytical methods*

7.4.1 Statistical model

This randomized clinical trial aims primarily to describe and estimate efficacy parameters and test pre-specified statistical hypotheses.

The primary variable will be analyzed using a logistic regression model, including, in addition to the treatment group, the stratification factor used at randomization (study site, previous enrollment in ARCTIC). Because we want to assess the treatment effect across centers, centre will be included as a covariate in the analysis. The statistical analysis plan (SAP) will detail these procedures, as well as alternative and further supportive evaluations, such as analyses including unbalanced baseline predictors or modifications of the logistic regression model in case validity assumptions are not met.

The primary analysis will be performed separately in patients on stable TNFi treatment (Group B) and in patients on stable synthetic DMARD treatment (Group S) of the primary per protocol population.

*7.4.1.1 Primary variable*

**Statistical hypothesis (non-inferior test)**

Null hypotheses:

- The difference in proportion of patients who are non-failures during the first period when comparing continued stable TNFi treatment (B1) to step-down and withdrawal of TNFi (B0) is at least 20% in favor of continued treatment.

- The difference in proportion of patients who are non-failures during the first period when comparing continued stable synthetic DMARD treatment (S1) to step-down of synthetic DMARDs (S1/2) is at least 20% in favor of continued treatment.

Alternative hypotheses:

- The difference in proportion of patients who are non-failures during the first period is less than 20% between continued full dose TNFi treatment (B1) and step-down to withdrawal of TNFi (B0).
- The difference in proportion of patients who are non-failures during the first period is less than 20% between continued full dose synthetic DMARD treatment (S1) and step-down to half dose of synthetic DMARD treatment (S1/2).

The primary variable will be evaluated by the 95% confidence limits. A conclusion of non-inferiority will be made if the 95% confidence limits of the estimated treatment difference fully lie within the inferiority margin of 20%. The inferiority margin of 20% was chosen based on the belief that both clinicians and patients would accept this when keeping the high costs and potential adverse events of biologic DMARDs in mind.

If the data indicate that one of the groups is superior to the other, a superiority test on the 5% significance level will be performed in each randomized group separately according to the following hypothesis:

**Statistical hypothesis (superiority test):**

Null hypotheses:

- There is no difference between continued full dose of TNFi (B1) and step down to withdrawal of TNFi (B0) with regards to the proportion of patients who are non-failures during the first study period.
- There is no difference between continued full dose of synthetic DMARD (S1) and step down to step down to half dose of synthetic DMARD (S1/2) with regards to the proportion of patients who are non-failures during the first study period.

Alternate hypotheses:

- There is a difference between continued full dose of TNFi (B1) and step down to withdrawal of TNFi (B0) with regards to the proportion of patients who are non-failures during the first study period, i.e. a two-sided test will be employed.

- There is a difference between continued full dose of synthetic DMARD (S1) and step down to half dose of synthetic DMARD (S1/2) with regards to the proportion of patients who are non-failures during the first study period, i.e. a two-sided test will be employed.

These analyses will be performed separately in patients in the ITT population in group B and group S.

*7.4.1.2 Secondary/exploratory variables*

Exploratory between-group comparisons will be performed for the primary endpoint on secondary and tertiary populations in addition to secondary efficacy endpoints on all populations. No adjustment of significance level for multiple comparisons will be performed.

7.4.3 Hypotheses testing

The primary efficacy analyses will be performed using a logistic regression model (see section 7.2.1 above for details on statistical model).

The between-group comparisons for secondary variables will be tested as for the primary variable where applicable and additional analyses will be performed based on the following methods (but not limited to):

- Continuous secondary variables will be subject to repeated measures mixed models or appropriate non-parametric alternatives
- Binary response variables will be analyzed using logistic regression (possibly adjusting for within-subject dependencies by generalized estimating equations) or chi-square/Mantel-Haenszel test
- Time-to-event variables will be analyzed using the Kaplan-Meier method and comparisons between the two groups will be performed using the log rank test or Cox regression analyses.

Unless otherwise specified, all statistical hypotheses will be tested as the primary variable, i.e. with an assessment of non-inferiority based on the 95% confidence limits of the estimated difference between the groups. Non-inferiority margins of the secondary endpoints will be specified in the SAP prior to analysis. If the data indicate that one of the groups is superior to the other, a superiority test on the 5% significance level will be performed as for the primary variable on the ITT population. Thus, the non-inferiority analyses will be done in the PP population and the superiority analyses in the ITT population as appropriate.

7.4.4 Efficacy analyses

All patients included in the PP population will form the primary analysis population of the study. All efficacy analyses will be presented with the results from the hypothesis testing (by p-value) in addition to estimates and 95% confidence limits of the treatment effect. For the primary variable specifically, this will be the estimated odds ratio and relative risk with corresponding 95% confidence limits.

7.4.5 Safety analyses

The safety analyses population will include all patients who completed at least one follow-up visit. Safety analyses will be descriptive and presented as summary tables by treatment group and (if applicable) by visit.

7.4.6 Quality of life and disability analyses

Quality of life and disability will be assessed using SF-36, EQ-5D and HAQ. These scores will be summarized by descriptive summary tables at baseline and over time, and at the end of study. Missing data at end of study will be replaced by the last valid post-baseline assessment.

7.4.7 Other analyses/subanalyses

Exploratory subgroup analyses of primary, secondary and exploratory efficacy variables may be performed if appropriate. The decision to include such analyses will be made on basis of the collected data.

7.4.8 Missing data

In the primary per protocol population we expect few or no missing data. If, however, missing data is regarded as having a significant effect on the conclusions of the trial, sensitivity analyses with different methods for handling missing data will be included. Such methods may include complete case analyses, last observation carried forward, worst case/best case imputation and multiple imputation techniques.

### *7.5 Randomization method*

We will apply a block randomization scheme with stratification according to center, and whether the patient previously participated in the ARCTIC trial. The randomization list will be

generated prior to study start, and will provide information on randomized treatment both for the first and second period. For Group S, the randomized treatment for the second period will not be revealed until the fourth visit. Allocation to randomized treatment will be done by sequentially choosing the next treatment from the randomization list according to stratification. The allocation will be performed by the EDC system (Viedoc). Patients in the **ARCTIC follow-up group** will not be randomized, but evaluated at each visit and continue to receive treatment according to the predefined ARCTIC protocol. If the patient during the follow-up time fulfils inclusion criteria for Group B or Group S, he or she can be included into the correct group and randomized (see also Table1).

### *7.6 Health economic models*

Long-term consequences of the three management strategies (full dose of DMARDs, reduced dose of DMARDs, discontinuation of DMARDs) will be evaluated in a RA markov model (NORAM – Norwegian Rheumatoid Arthritis Model) that is currently under development (at the Department of Health Economics, UiO), and in other appropriate health economics models. The NORAM model captures health outcomes in terms of QALYS (Quality Adjusted Life Years) based on SF-36 data, and use of resources in monetary terms (pharmaceuticals, in-hospital and out-hospital care, rehabilitation, physiotherapy, indirect costs, etc.). The model will be based primarily on the trial data, but extrapolation in time will be performed based on other data sources.

### *7.7 Definitions of flare in RA and collaboration with the OMERACT flare group*

The data collection in ARCTIC REWIND will include several measures that could contribute to validation of flare definitions in RA, including the provisional OMERACT flare questionnaire. Dr. Elisabeth Lie will have the main responsibility for the collaboration with the OMERACT flare group and the scientific content of this objective.

## 8. Ethical aspects

The study will be conducted in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with ICH/Good Clinical Practice and applicable regulatory requirements. Registration of patient data will be carried out in accordance with national personal data laws.

The study will be reviewed by the local "personombud", and permission to store sensitive data will be obtained from the Norwegian Data Inspectorate. A biobank will be established according to Norwegian regulations. The study will be conducted in full conformance with the principles of the "Declaration of Helsinki". The study will adhere to the principles outlined in "Guideline for Good Clinical Practice" ICH Tripartite Guideline (January 1997).

### *8.1 Ethics Committee Approval*

The study protocol, including the patient information and informed consent form to be used, must be approved by the regional ethics committee before enrolment of any patients into the study. The investigator is responsible for informing the ethics committee of any serious and unexpected adverse events and/or major amendments to the protocol as per national requirements.

### *8.2 Other Regulatory Approvals*

The protocol will be submitted and approved by the applicable competent authorities before commencement of the study. The protocol will also be registered in www.clinicaltrials.gov before inclusion of the first patient.

### *8.3 Informed Consent*

The investigator is responsible for giving the patients full and adequate verbal and written information about the nature, purpose, possible risk and benefit of the study. They will be informed as to the strict confidentiality of their patient data, but that their medical records may be reviewed for trial purposes by authorized individuals other than their treating physician.
It will be emphasized that the participation is voluntary and that the patient is allowed to refuse further participation in the protocol whenever she/he wants. This will not prejudice the patient's subsequent care. Documented informed consent must be obtained for all patients included in

the study before they are registered in the study. This will be done in accordance with the national and local regulatory requirements. The investigator is responsible for obtaining signed informed consent.

A copy of the patient information and consent will be given to the patients. The signed and dated patient consent forms will be filed in the Investigator File binder.
It is the responsibility of the treating rheumatologist (or a person designated by him/her, i.e. a study nurse) to obtain signed informed consent from each patient prior to participating in this study after adequate explanation of the aims, methods, anticipated benefits and potential hazards of the study.

### *8.4 Subject Identification*

The investigator is responsible for keeping a list of all patients (who have received study treatment or undergone any study specific procedure) including patient numbers, full names and last known addresses. The patients will be identified in the CRFs by patient number, initials and date of birth.

## 9. Source documents and case report form completion

Adequate and accurate records of each study visit will be maintained to enable the conduct of the study to be fully documented and the study data to be subsequently verified.

### *9.1 Source documents*

The medical records for each patient should contain information which is important for the patient's safety and continued care and to fulfil the requirement that critical study data should be verifiable. To achieve this, the medical records of each patient should clearly describe at least:

- That the patient is participating in the study, e.g. by including the enrollment number and the study code or other study identification;
- Date when Informed Consent was obtained from the patient and statement that patient received a copy of the signed and dated Informed Consent;
- Results of all assessments confirming a patient's eligibility for the study;
- Treatments given, changes in treatments during the study and the time points for the changes;
- Visits to the clinic / telephone contacts during the study, including those for study purposes only;
- Date of, and reason for, discontinuation from study treatment;
- Date of, and reason for, withdrawal from study;
- Date of death and cause of death, if available;
- Additional information according to local regulations and practice.

### *9.2 Case report forms*

Accurate and reliable data collection will be assured by verification and cross-checking of the CRF/eCRF by the study nurse, study physician and external monitor. A web-based eCRF software solution that adheres to GCP will be used to collect study data (Viedoc™, Uppsala, Sweden). The GoTreatIT software package (DiagraphIT, Kristiansand, Norway), which is defined as being part of the electronic patient journal (EPJ), may be used as source document for the eCRF. If a centre does not have access to the GoTreatIT software package, or the GoTreatIT program do not offer the necessary specifications, paper files/paper CRF combined with the patient's EPJ will be used as source documents for the eCRF when applicable. Semi-

automated, validated procedures for importing non-identifiable information from the patient's EPJ (such as patient reported outcomes) into the eCRF may be generated and used.

### *9.3 Source data verification*

The investigator will be visited on a regular basis by the Clinical Study Monitor, who will check completed CRFs, discuss the progress of the study and monitor drug usage according to European ICH GCP. The monitoring will also include source data verification (SDV).

Sponsor's representatives (e.g. monitors, auditors) and/or regulatory authorities will be allowed access to source data for source data verification in which case a review of those parts of the hospital records relevant to the study may be required. This study will be monitored by the "Department of Clinical Research Support, Oslo University Hospital" and "Innovest AS". Monitoring is the act of overseeing the progress of a clinical trial, and of ensuring that it is conducted, recorded, and reported in accordance with the protocol, Standard Operating Procedures (SOPs), Good Clinical Practice (GCP), and the applicable regulatory requirement(s). It is understood that the responsible monitor will contact and visit the trial sites regularly, and will be allowed, on request, to inspect the various records of the trial (eCRF and other records/source material), provided that patient confidentiality is maintained in accord with local requirements.

### *9.4 Storage of study documentation*

The investigator shall arrange for the retention of the patient identification and the code list. Patient files shall be kept for the maximum period of time permitted by each hospital. The study documentation (CRFs, Site File etc) shall be retained and stored during the study and for 15 years after study closure. All information concerning the study will be stored in a safe place inaccessible to unauthorized personnel.

## 10. Study committee

A study committee will be established, and will primarily review safety data as needed, as well as make recommendations regarding continuation, termination or modification of the study.

The study committee will include the national medical leader, the principal investigator, the biostatistician, the project leader, an internationally recognized medical professor/ epidemiologist, a local investigator, a patient representative and a representative from Diakonhjemmet Sykehus AS.

## 11. Publication policy

The results of this study will be published and/or presented at scientific meetings and in international peer-reviewed journals. Authorship will be based on scientific contribution and enrolment, according to the guidelines set forth in the Vancouver protocol.

## 12. Financial aspects and insurance

The medical treatment will be covered as for "usual care", i.e. by "Folketrygden"/"NAV" and patients own payment "egenandel". Procedures/examinations that are not part of "usual care" will be covered by the study, and there will be no own payments.

All patients will receive a reimbursement of NOK 100 per visit, to cover parking fees as the study implies longer visits than "usual care".

All centers will receive a reimbursement of NOK 500 per visit (with payments after final visit), to cover extra resources allocated to the study, including study nurses and administration of the logistics of the study.

Diakonhjemmet Sykehus AS and participating centers will cover necessary insurance in accordance with Norwegian law (through "Legemiddelansvarsforeningen – LAF").

Funding will partly be from Helse Sør-Øst research grant to Diakonhjemmet Hospital (Helse Sør-Øst research Group in rheumatology) and other public funding schemes, partly by grants

from pharmaceutical companies and support from the research foundation of the Norwegian Rheumatism Association or other non-profit organizations.

## 13. Study management

### *13.1 Investigator Delegation Procedure*

The principal investigator is responsible for making and updating a "delegation of tasks" listing all the involved co-workers and their role in the project. He will ensure that appropriate training relevant to the study is given to all of these staff, and that any new information of relevance to the performance of this study is forwarded to the staff involved.

### *13.2 Study Amendments*

If it is necessary for the study protocol to be amended, the amendment and/or a new version of the study protocol (Amended Protocol) must be notified to and approved by the Competent Authority and the Ethics Committee according to EU and national regulations.

### *13.3 Audit and Inspections*

Authorized representatives of a regulatory authority and Ethics Committee may visit the centre to perform inspections, including source data verification. Likewise the representatives from sponsor may visit the center to perform an audit. The purpose of an audit or inspection is to systematically and independently examine all study-related activities and documents to determine whether these activities were conducted, and data were recorded, analyzed, and accurately reported according to the protocol, Good Clinical Practice (ICH GCP), and any applicable regulatory requirements. The principal investigator will ensure that the inspectors and auditors will be provided with access to source data/documents.

## 14. References

(1) Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet 2010 Sep 25;376(9746):1094-108.

(2) Smolen JS, Landewe R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis 2010 Jun;69(6):964-75.

(3) Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken ) 2012 May;64(5):625-39.

(4) Combe B, Landewe R, Lukas C, Bolosiu HD, Breedveld F, Dougados M, et al. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2007 Jan;66(1):34-45.

(5) Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010 Apr;69(4):631-7.

(6) Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. Arthritis Rheum 2003 Jul;48(7):1771-4.

(7) SHORT CL, BAUER W. The course of rheumatoid arthritis in patients receiving simple medical and orthopedic measures. N Engl J Med 1948 Jan 29;238(5):142-8.

(8) Pinals RS, Masi AT, Larsen RA. Preliminary criteria for clinical remission in rheumatoid arthritis. Arthritis Rheum 1981 Oct;24(10):1308-15.

(9) Prevoo ML, van Gestel AM, van 'T Hof M, van Rijswijk MH, van de Putte LB, van Riel PL. Remission in a prospective study of patients with rheumatoid arthritis. American Rheumatism Association preliminary remission criteria in relation to the disease activity score. Br J Rheumatol 1996 Nov;35(11):1101-5.

(10) van Riel PL, van Gestel AM. Clinical outcome measures in rheumatoid arthritis. Ann Rheum Dis 2000 Nov;59 Suppl 1:i28-i31.

(11) Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol 2005 Sep;23(5 Suppl 39):S100-S108.

(12) Boers M, Felson DT, Wells G, van Tuyl LH, Zhang B, Funovits J, et al. Progress toward the development of a new definition of remission in rheumatoid arthritis. Bull NYU Hosp Jt Dis 2010;68(2):140-2.

(13) Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American college of rheumatology/european league against rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Ann Rheum Dis 2011 Mar;70(3):404-13.

(14) Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum 2011 Mar;63(3):573-86.

(15) van Tuyl LH, Vlad SC, Felson DT, Wells G, Boers M. Defining remission in rheumatoid arthritis: results of an initial American College of Rheumatology/European League Against Rheumatism consensus conference. Arthritis Rheum 2009 May 15;61(5):704-10.

(16) Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. Arthritis Rheum 2008 Oct;58(10):2958-67.

(17) Scire CA, Montecucco C, Codullo V, Epis O, Todoerti M, Caporali R. Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. Rheumatology (Oxford) 2009 Sep;48(9):1092-7.

(18) Peluso G, Michelutti A, Bosello S, Gremese E, Tolusso B, Ferraccioli G. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. Ann Rheum Dis 2011 Jan;70(1):172-5.

(19) Gandjbakhch F, Foltz V, Mallet A, Bourgeois P, Fautrel B. Bone marrow oedema predicts structural progression in a 1-year follow-up of 85 patients with RA in remission or with low disease activity with low-field MRI. Ann Rheum Dis 2011 Dec;70(12):2159-62.

(20) Saleem B, Brown AK, Quinn M, Karim Z, Hensor EM, Conaghan P, et al. Can flare be predicted in DMARD treated RA patients in remission, and is it important? A cohort study. Ann Rheum Dis 2012 Aug;71(8):1316-21.

(21) Foltz V, Gandjbakhch F, Etchepare F, Rosenberg C, Tanguy ML, Rozenberg S, et al. Power Doppler ultrasound, but not low-field magnetic resonance imaging, predicts relapse and radiographic disease progression in rheumatoid arthritis patients with low levels of disease activity. Arthritis Rheum 2012 Jan;64(1):67-76.

(22) Yoshimi R, Hama M, Takase K, Ihata A, Kishimoto D, Terauchi K, et al. Ultrasonography is a potent tool for the prediction of progressive joint destruction during clinical remission of rheumatoid arthritis. Mod Rheumatol 2012 Jul 18.

(23) Jacobsson LT, Hetland ML. New remission criteria for RA: 'modern times' in rheumatology--not a silent film, rather a 3D movie. Ann Rheum Dis 2011 Mar;70(3):401-3.

(24) Johnson K. A suggestion and the larger perspective on the American College of Rheumatology/European League Against Rheumatism remission definitions: comment on the article by Felson et al. Arthritis Rheum 2011 Nov;63(11):3643-4.

(25) van der Heijde D. Remission by imaging in rheumatoid arthritis: should this be the ultimate goal? Ann Rheum Dis 2012 Apr;71 Suppl 2:i89-92.:i89-i92.

(26) Felson D. Defining remission in rheumatoid arthritis. Ann Rheum Dis 2012 Apr;71 Suppl 2:i86-8.:i86-i88.

(27) van den BM, Klarenbeek NB, Dirven L, van SD, Hulsmans HM, Kerstens PJ, et al. Discontinuation of infliximab and potential predictors of persistent low disease activity in patients with early rheumatoid arthritis and disease activity score-steered therapy: subanalysis of the BeSt study. Ann Rheum Dis 2011 Aug;70(8):1389-94.

(28) Tanaka Y, Takeuchi T, Mimori T, Saito K, Nawata M, Kameda H, et al. Discontinuation of infliximab after attaining low disease activity in patients with rheumatoid arthritis: RRR (remission induction by Remicade in RA) study. Ann Rheum Dis 2010 Jul;69(7):1286-91.

(29) Brocq O, Millasseau E, Albert C, Grisot C, Flory P, Roux CH, et al. Effect of discontinuing TNFalpha antagonist therapy in patients with remission of rheumatoid arthritis. Joint Bone Spine 2009 Jul;76(4):350-5.

(30) Kavanaugh A. Withdrawal of Adalimumab in Early Rheumatoid Arthritis Patients Who Attained Stable Low Disease Activity with Adalimumab Plus Methotrexate: Results of a Phase 4, Double-Blind, Placebo-Controlled Trial. ACR 2011.

(31) Detert J, Bastian H, Listing J, Weiss A, Wassenberg S, Liebhaber A, et al. Induction therapy with adalimumab plus methotrexate for 24 weeks followed by methotrexate monotherapy up to week 48 versus methotrexate therapy alone for DMARD-naive patients with early rheumatoid arthritis: HIT HARD, an investigator-initiated study. Ann Rheum Dis 2012 Jul 10.

(32) van VR. In Rheumatoid Arthritis Patients with Stable Low Disease Activity On Methotrexate Plus Etanercept, Continuation of Etanercept 50 Mg Weekly or 25 Mg Weekly Are Both Clinically Superior to Discontinuation: Results From a Randomized, 3-Armed, Double-Blind Clinical Trial. ACR 2012 2012.

(33) Tishler M, Caspi D, Yaron M. Methotrexate treatment of rheumatoid arthritis: is a fortnightly maintenance schedule enough? Ann Rheum Dis 1992 Dec;51(12):1330-1.

(34) Kremer JM, Rynes RI, Bartholomew LE. Severe flare of rheumatoid arthritis after discontinuation of long-term methotrexate therapy. Double-blind study. Am J Med 1987 Apr;82(4):781-6.

(35) Cade R, Stein G, Pickering M, Schlein E, Spooner G. Low dose, long-term treatment of rheumatoid arthritis with azathioprine. South Med J 1976 Apr;69(4):388-92.

(36) ten WS, Breedveld FC, Hermans J, Vandenbroucke JP, van de Laar MA, Markusse HM, et al. Randomised placebo-controlled study of stopping second-line drugs in rheumatoid arthritis. Lancet 1996 Feb 10;347(8998):347-52.

(37) Ahern MJ, Hall ND, Case K, Maddison PJ. D-penicillamine withdrawal in rheumatoid arthritis. Ann Rheum Dis 1984 Apr;43(2):213-7.

(38) Klarenbeek NB, van der Kooij SM, Guler-Yuksel M, van Groenendael JH, Han KH, Kerstens PJ, et al. Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. Ann Rheum Dis 2011 Feb;70(2):315-9.

(39) van der WD, Visser K, Klarenbeek NB, Ronday HK, Peeters AJ, Kerstens PJ, et al. Sustained drug-free remission in rheumatoid arthritis after DAS-driven or non-DAS-driven therapy: a comparison of two cohort studies. Rheumatology (Oxford) 2012 Jun;51(6):1120-8.

(40) Rantalaiho V, Korpela M, Hannonen P, Kautiainen H, Jarvenpaa S, Leirisalo-Repo M, et al. The good initial response to therapy with a combination of traditional disease-modifying antirheumatic drugs is sustained over time: the eleven-year results of the Finnish rheumatoid arthritis combination therapy trial. Arthritis Rheum 2009 May;60(5):1222-31.

(41) Machold KP. Prevention and cure of rheumatoid arthritis: is it possible? Best Pract Res Clin Rheumatol 2010 Jun;24(3):353-61.

(42) Tanaka Y. Intensive treatment and treatment holiday of TNF-inhibitors in rheumatoid arthritis. Curr Opin Rheumatol 2012 May;24(3):319-26.

(43) Goekoop-Ruiterman YP, Huizinga TW. Rheumatoid arthritis: can we achieve true drug-free remission in patients with RA? Nat Rev Rheumatol 2010 Feb;6(2):68-70.

(44) O'Mahony R, Richards A, Deighton C, Scott D. Withdrawal of disease-modifying antirheumatic drugs in patients with rheumatoid arthritis: a systematic review and meta-analysis. Ann Rheum Dis 2010 Oct;69(10):1823-6.

(45) Krieckaert C, Rispens T, Wolbink G. Immunogenicity of biological therapeutics: from assay to patient. Curr Opin Rheumatol 2012 May;24(3):306-11.

(46) Krieckaert CL, Bartelds GM, Lems WF, Wolbink GJ. The effect of immunomodulators on the immunogenicity of TNF-blocking therapeutic monoclonal antibodies: a review. Arthritis Res Ther 2010;12(5):217.

(47) Krieckaert CL, Bartelds GM, Wolbink GJ. Therapy: Immunogenicity of biologic therapies-we need tolerance. Nat Rev Rheumatol 2010 Oct;6(10):558-9.

(48) Krieckaert CL, Nurmohamed MT, Wolbink GJ. Methotrexate reduces immunogenicity in adalimumab treated rheumatoid arthritis patients in a dose dependent manner. Ann Rheum Dis 2012 Nov;71(11):1914-5.

(49) Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010 Apr;69(4):631-7.

(50) Kvamme MK, Lie E, Kvien TK, Kristiansen IS. Two-year direct and indirect costs for patients with inflammatory rheumatic joint diseases: data from real-life follow-up of

patients in the NOR-DMARD registry. Rheumatology (Oxford) 2012 Sep;51(9):1618-27.

(51) Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010 Sep;69(9):1580-8.

(52) Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, III, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010 Sep;69(9):1580-8.

(53) van der Helm-van Mil AH, Breedveld FC, Huizinga TW. Aspects of early arthritis. Definition of disease states in early arthritis: remission versus minimal disease activity. Arthritis Res Ther 2006;8(4):216.

(54) van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol 1999 Mar;26(3):743-5.

(55) Hammer HB, Bolton-King P, Bakkeheim V, Berg TH, Sundt E, Kongtorp AK, et al. Examination of intra and interrater reliability with a new ultrasonographic reference atlas for scoring of synovitis in patients with rheumatoid arthritis. Ann Rheum Dis 2011 Nov;70(11):1995-8.

(56) Ostergaard M, Peterfy C, Conaghan P, McQueen F, Bird P, Ejbjerg B, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. J Rheumatol 2003 Jun;30(6):1385-6.

(57) Haavardsholm EA, Ostergaard M, Ejbjerg BJ, Kvan NP, Kvien TK. Introduction of a novel magnetic resonance imaging tenosynovitis score for rheumatoid arthritis: reliability in a multireader longitudinal study. Ann Rheum Dis 2007 Sep;66(9):1216-20.

(58) Ostergaard M, Boyesen P, Eshed I, Gandjbakhch F, Lillegraven S, Bird P, et al. Development and preliminary validation of a magnetic resonance imaging joint space narrowing score for use in rheumatoid arthritis: potential adjunct to the OMERACT RA MRI scoring system. J Rheumatol 2011 Sep;38(9):2045-50.

(59) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-83.

(60) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002 Mar;21(2):271-92.

(61) Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982 Sep;9(5):789-93.

(62) Lillegraven S, Kvien TK. Measuring disability and quality of life in established rheumatoid arthritis. Best Pract Res Clin Rheumatol 2007 Oct;21(5):827-40.

(63) Brooks R. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72.

(64) EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990 Dec;16(3):199-208.

(65) Gossec L, Paternotte S, Aanerud GJ, Balanescu A, Boumpas DT, Carmona L, et al. Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient-derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. Ann Rheum Dis 2011 Jun;70(6):935-42.

(66) Kvien TK, Heiberg, Lie E, Kaufmann C, Mikkelsen K, Nordvag BY, et al. A Norwegian DMARD register: prescriptions of DMARDs and biological agents to patients with inflammatory rheumatic diseases. Clin Exp Rheumatol 2005 Sep;23(5 Suppl 39):S188-S194.

## 15. Appendix

### *15.1 Appendix 1 Overview of visits*

Before or at the baseline visit the following screening procedure is performed:
- Physical examination including pulse rate and blood pressure (systolic and diastolic)
- Laboratory tests (CRP, ESR, hematology, blood chemistry,urine dipstick)
- Cross-check of inclusion- and exclusion criteria

**Table:** Data collection and timing of each ARCTIC REWIND visit

| Visits | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Flare visit* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Months | 0 | 4 | 8 | 12 | 16 | 20 | 24 | 28 | 32 | 36 | |
| **History** | | | | | | | | | | | |
| Informed consent / Medical history / Eligibility assessment | X | | | | | | | | | | |
| Comorbidities / Medication / Adverse events | X | X | X | X | X | X | X | X | X | X | X |
| Work performance (WPAI instrument) | X | X | X | X | X | X | X | X | X | X | X |
| Smoking / Coffee / Alcohol | X | | | X | | | X | | | X | |
| **Clinical assessments by nurse and/or physician** | | | | | | | | | | | |
| Assessment of disease activity, swollen and tender joint counts | X | X | X | X | X | X | X | X | X | X | X |
| Vital signs | X | X | X | X | X | X | X | X | X | X | X |
| **Patient reported variables** | | | | | | | | | | | |
| Pain VAS/Fatigue VAS/ Patients assessment of disease activity VAS | X | X | X | X | X | X | X | X | X | X | X |
| HAQ-PROMIS / EQ-5D / RAID / Flare questionnaire | X | X | X | X | X | X | X | X | X | X | X |
| SF-36 | X | | | X | | | X | | | X | X |
| **Biochemical examinations** | | | | | | | | | | | |
| ACPA / IgA-RF / IgM-RF | X | | | | | | | | | | |
| ESR / CRP / Hematology / ALT / Creatinine / GFR / Blood chemistry | X | X | X | X | X | X | X | X | X | X | X |
| BioBank (whole blood, serum, plasma, urine) | X | | | X | | | X | | | X | |
| **Radiologic assessments** | | | | | | | | | | | |
| MRI of dominant hand | X | | | X | | | X | | | X*** | |
| Radiographs of hands and feet | X | | | X | | | X | | | X | |
| Ultrasonography of 32 joints | X | ** | ** | X | ** | ** | X | ** | ** | X | X |

* If a patient suspects a disease activity flare, an extra visit should be scheduled within a week.
** If the patient has a flare in disease activity, ultrasonography should also be performed at the visit.
*** MRI at 36 months will be performed without gadolinium-based contrast agent.

### *15.2 Appendix 2 ACR/EULAR Classification Criteria for Rheumatoid Arthritis*

ACR/EULAR Classification Criteria for Rheumatoid Arthritis: score based algorithm for classification in an eligible patient (i.e. clinical synovitis not related to a specific aetiology). **Cutpoint for RA: ≥6/10**.

**Table 4.** ACR/EULAR Classification Criteria for Rheumatoid Arthritis

| | |
|---|---|
| JOINT INVOLVEMENT[1] (0-5) | |
| 1 medium-large[2] joint | 0 |
| 2-10 medium-large joints | 1 |
| 1-3 small[3] joints (with or without involvement of large joints) | 2 |
| 4-10 small joints (with or without involvement of large joints) | 3 |
| >10 joints[4] (at least one small joint) | 5 |
| *SEROLOGY[5] (0-3) | |
| Negative RF AND negative ACPA | 0 |
| Low positive RF OR low positive ACPA | 2 |
| High positive RF OR high positive ACPA | 3 |
| *ACUTE PHASE REACTANTS[6] (0-1) | |
| Normal CRP AND normal ESR | 0 |
| Abnormal CRP OR abnormal ESR | 1 |
| DURATION OF SYMPTOMS[7] (0-1) | |
| <6 weeks | 0 |
| ≥6 weeks | 1 |

[1] Joint involvement refers to any *swollen* or *tender* joint on examination, or evidence of synovitis on magnetic resonance imaging or ultrosonography. Distal interphalangeal joints (DIPs), 1st carpo-metacarpal (CMC) joint, and 1st metatarso-phalangeal (MTP) joint are ***excluded from assessment.*** Categories of joint distribution are classified according to the location and number of the involved joints, with placement into the highest category possible based on the pattern of joint involvement

[2] Medium to large joints refer to shoulders, elbows, hips, knees, and ankles

[3] Small joints refer to the metacarpo-phalangeal (MCP) joints, proximal interphalangeal (PIP) joints, metatarso-phalangeal (MTP) joints 2-5, thumb interphalangeal (IP) joints, and wrists

[4] In this category, at least one of the involved joints must be a small joint; the other joints can include any combination of large and additional small joints, as well as other joints not specifically listed elsewhere (*e.g.*, temporomandibular, acromioclavicular, sternoclavicular, *etc.*)

[5] Negative refers to international unit (IU) values that are ≤upper limit of normal (ULN) for the lab and assay; low-positive refers to IU values that are >ULN but ≤3x ULN for the lab and assay; high-positive referes to IU values that are >3x ULN for the lab and assay. Where RF is only available as positive or negative, a positive result should be scored as "low-positive" for RF.

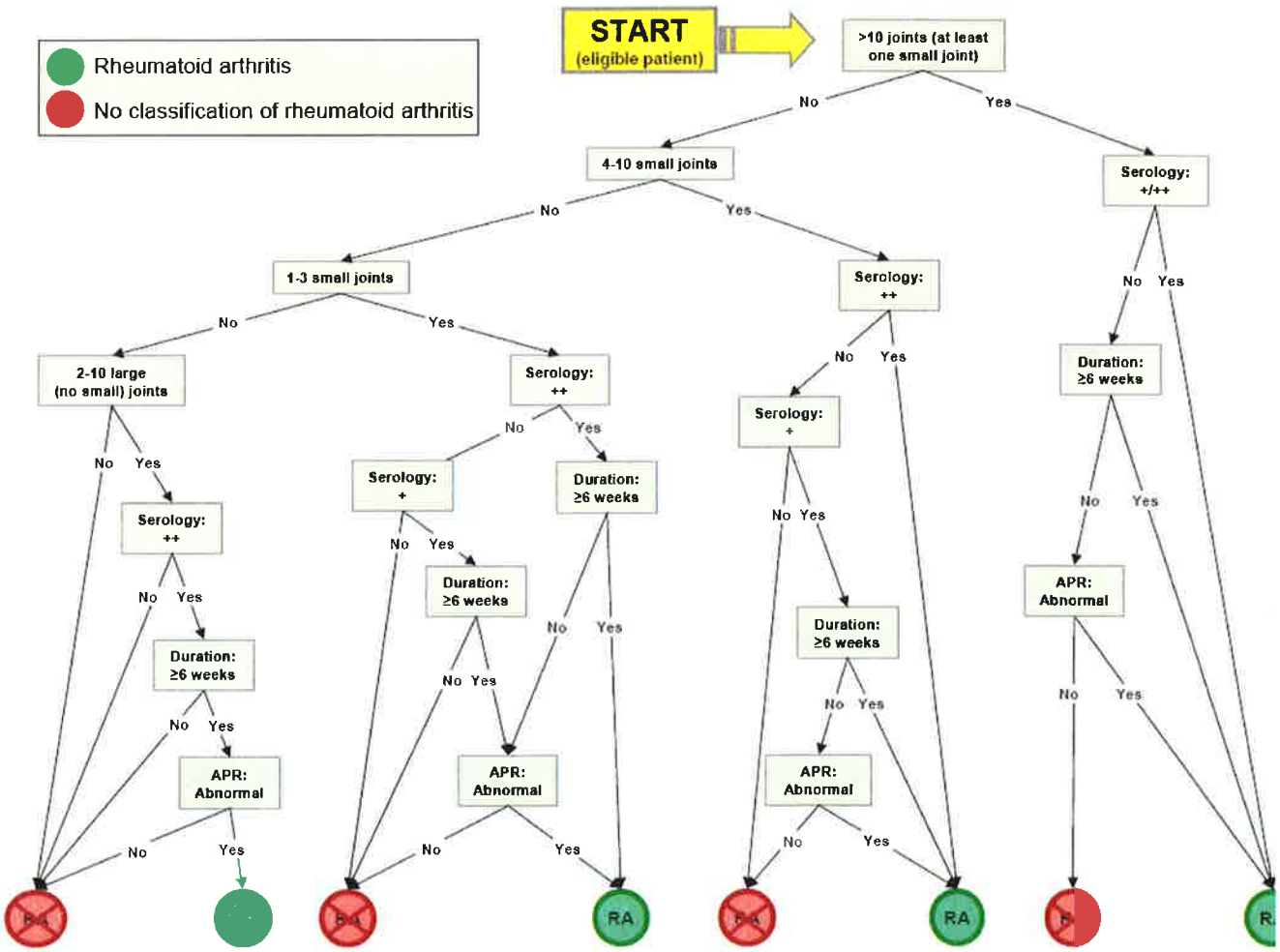
[6] Normal/abnormal is determined by local laboratory standards.

[7] Duration of symptoms refers to patient self-report of the duration of signs or symptoms of synovitis (*e.g.*, pain, swelling, tenderness) of joints that are clinically involved at the time of assessment.

* Individuals should only be scored by these criteria if at least one serologic test and at least one acute phase reactant test result is available. Where a value for a serologic test or acute phase reactant is not available, that test should be considered as 'negative/normal'.

**Abbreviations:** RF = rheumatoid factor; ACPA = anti-citrullinated protein/peptide antibodies; ULN = upper limit of normal; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein.

**Figure 1.** Tree algorithm to classify definite rheumatoid arthritis (green circles) or to exclude its current presence (red circles) among those who are eligible to be assessed by the new criteria. For definitions of categories (e.g. serology: +, ++, or joint regions) please see footnotes to Table 4.



### *15.3 Appendix 3 ACR core set*

1. Patient's evaluation of pain
2. Patient's global assessment of disease status
3. Physician's global assessment of disease status
4. Swollen joints
5. Tender joints
6. ESR/CRP
7. Health Assessment Questionnaire (HAQ)

## 15.4 Appendix 4 Joints to assessed for swelling and tenderness

**Ritchie Articular Index**
This is a long-standing approach to doing a graded assessment of the tenderness of 26 joint regions, based on summation of joint responses after applying firm digital pressure.24 Four grades can be used: 0, patient reported no tenderness; +1, patient complained of pain; +2, patient complained of pain and winced; and +3, patient complained of pain, winced, and withdrew. Thus, the index ranges from 0 to 3 for individual measures and 0 to 78 overall, with higher scores being worse tenderness.

Certain joints are treated as a single unit, such as the metacarpal-phalangeal and proximal interphalangeal joints of each hand and the metatarsal-phalangeal joints of each foot. For example, the maximum score for the five metacarpal-phalangeal joints of the right hand would be 3, not 15.

**44 Swollen joint count**
A 44 swollen joint count (see table below) is included in the DAS and includes the sternoclavicular and acromioclavicular joints, the shoulders, elbows, wrists, knees, ankles, and MCP, PIP, and MTP joints.

**Figure 2**, overview of joint assessments, from Sokka and Pincus, Clin Exp Rheumatol 2005; 23 (Suppl. 39):S58-S62.

| Joint | 66 68 joints (15) | Ritchie Index (16) | 44 joints (22) | 6 joints (23) | 28 joints (19) | 42 joints (24) |
|---|---|---|---|---|---|---|
| Temporomandibular | + | –* | | | | |
| Sternoclavicular | + | –* | – | | | |
| Acromioclavicular | + | –* | – | | | |
| Shoulder | + | – | – | | – | – |
| Elbow | – | – | + | – | – | – |
| Wrist | + | + | – | – | – | – |
| Metacarpophalangeal | | – | | | | |
| First | + | | – | | – | – |
| Second | + | | – | – | – | – |
| Third | + | | – | – | – | – |
| Fourth | – | | + | – | – | – |
| Fifth | + | | – | – | – | – |
| Proximal interphalangeal | | + | | | | |
| First | – | | + | – | – | – |
| Second | + | | – | – | – | – |
| Third | + | | – | + | – | – |
| Fourth | + | | – | – | – | – |
| Fifth | + | | + | + | – | – |
| Distal interphalangeal | | | | | | |
| Second | – | | | | | |
| Third | + | | | | | |
| Fourth | + | | | | | |
| Fifth | + | | | | | |
| Hip | +# | + | | | | – |
| Knee | + | + | – | – | – | – |
| Ankle | – | – | + | – | | – |
| Talocalcaneal | | + | | | | |
| Tarsus | + | + | | | | |
| Metatarsophalangeal | | + | | | | |
| First | + | | – | – | | – |
| Second | – | | + | – | | – |
| Third | + | | – | – | | – |
| Fourth | + | | – | – | | – |
| Fifth | + | | – | – | | – |
| Proximal interphalangeal (toe) | | | | | | |
| First | + | | | | | |
| Second | + | | | | | |
| Third | + | | | | | |
| Fourth | + | | | | | |
| Fifth | + | | | | | |

#Assessed for tenderness only. *right and left joints assessed together.

### *15.5 Appendix 5 SF-36*

The SF-36 is a multi-purpose, short-form health survey with 36 questions. It yields an 8-scale profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures and a preference-based health utility index. It is a generic measure, as opposed to one that targets a specific age, disease, or treatment group. Accordingly, the SF-36 has proven useful in surveys of general and specific populations, comparing the relative burden of diseases, and in differentiating the health benefits produced by a wide range of different treatments.

See CRF for the wording of the questionnaire.

### *15.6 Appendix 6 EQ-5D*

See CRF for the wording of the questionnaire.

### *15.7 Appendix 7 HAQ-promis*

The Health Assessment Questionnaire (HAQ) was originally developed in 1978 by James F. Fries, MD, and colleagues at Stanford University. It was one of the first self-report functional status (disability) measures and has become the dominant instrument in many disease areas, including arthritis. It is widely used throughout the world and has become a mandated outcome measure for clinical trials in rheumatoid arthritis and some other diseases. In this trial we will apply the updated version called HAQ-promis, which has one extra response category compared to the original HAQ.

Patients usually find the instruments self-explanatory. Clarification is seldom required.

**PROMIS 20-item Physical Function Short Form**

Response pattern: Without any difficulty - With a little difficulty - With some difficulty - With much difficulty - Unable to do

Are you able to do chores such as vacuuming or yard work?
Are you able to run a short distance, such as to catch a bus?
Are you able to get in and out of a car?
Are you able to push open a heavy door?
Are you able to dry your back with a towel?
Are you able to change a light bulb overhead?
Are you able to hold a plate full of food?
Are you able to transfer from a bed to chair and back?
Are you able to dress yourself, including tying shoelaces and doing buttons?
Are you able to wash and dry your body?
Are you able to shampoo your hair?
Are you able to squeeze a new tube of toothpaste?
Are you able to sit on the edge of a bed?
Are you able to get on and off the toilet?

Response pattern: Not at all - Very little - Somewhat - Quite a lot - Cannot do

Does your health now limit you in lifting or carrying groceries?
Does your health now limit you in doing two hours of physical labor?
Does your health now limit you in doing vigorous activities, such as running, lifting heavy objects, participating in strenuous sports?
Does your health now limit you in climbing one flight of stairs?
Does your health now limit you in bending, kneeling, or stooping?
Does your health now limit you in walking more than a mile?

### *15.8 Appendix 8 RAID*

See CRF for the wording of the questionnaire.

### *15.9 Appendix 9 EULAR response*

Based on the DAS, response criteria have been developed: the EULAR response criteria. The EULAR response criteria include not only change in disease activity but also current disease activity. To be classified as responders, patients should have a significant change in DAS and also low current disease activity. Three categories are defined: good, moderate, and non-responders.

### *15.10 Appendix 10 Sharp van der Heijde Score*

Radiographs of hands (posterior/anterior) and foot (anterior/posterior) will be taken at baseline, 3, 6, 12, 16 and 24 months. The modified Sharp van der Heijde Score (vdHSS) will be calculated, including an erosion score and a joint space narrowing score.

According to the vdHSS, erosion is assessed in 16 joints (five MCP, four PIP, IP of the thumbs, 1st MCB, radius and ulna bones, trapezium and trapezoid as one unit (multangular), navicular, lunate) for each hand and wrist, and six joints (five MTP, IP) for each foot. One point is scored if erosions are discrete, rising to 2, 3, 4, or 5 depending on the amount of surface area affected. The score for erosion ranges from 0 to 160 in the hands and from 0 to 120 in the feet (the maximum erosion score for a joint in the foot is 10). JSN is assessed in 15 joints (five MCP, four PIP, CMC 3 to 5, multangular navicular-lunate, radiocarpal) for each hand and wrist, and six joints (five MTP, IP) for each foot. JSN is combined with a score for (sub)luxation and scored as follows: 0 = normal; 1 = focal or doubtful; 2 = generalised, less than 50% of the original joint space; 3 = generalised, more than 50% of the original joint space or subluxation; 4 = bony ankylosis or complete luxation. The score for JSN ranges from 0 to 120 in the hands and from 0 to 48 in the feet (50). The total score is the sum of scores of erosion and JSN, the maximum score being 448.

Original radiographs will be scored at a central reading site by two experienced readers.

### *15.11 Appendix 11 Ultrasonography score*

A Siemens Antares or a machine with similar qualities will be used for the ultrasonographic assessments, with the following (preliminary) specifications:
PRF = 391
Frequency = 7.3 MHz
R/S = 5
Filter = 2

The following joints will be scored for grey scale synovitis and power Doppler signal, with the best possible match to a reference atlas developed by Dr. Hilde Berner Hammer (accessible within the GoTreatIT software). The scoring system is based on the OMERACT US recommendations, with a semi-quantitative score from 0 - 3 in the following 32 joints: bilateral MCP I-V, RCJ, DRUJ, intercarpal, elbow, knee, talocrural joints and MTP I-V, yielding a maximum total score of 96 for both grey scale and power Doppler.

In addition to the following 32 joints the following will be scored: Bilateral PIP2 and 3 joints, extensor carpi ulnaris tendon and tibialis posterior tendon.

### *15.12 Appendix 12 RAMRIS*

Magnetic resonance imaging of the dominant hand and wrist with and without gadolinium will be undertaken at baseline, 12 and 24 months. At 36 months MRI will be performed without gadolinium-based contrast agent. The same side will be assessed at all time-points.

Images will be read according to the RAMRIS score at a central reading site by two experienced readers. The RAMRIS consists of MRI definitions of important joint pathologies, a core set of MRI sequences and a semi-quantitative scoring system for erosions, bone marrow edema and synovitis (59). The RAMRIS core set of MRI sequences to assess inflammatory as well as destructive changes in RA joints includes: Imaging in 2 planes, with T1-weighted images before and after intravenous gadolinium-contrast to assess synovitis and erosions; plus a T2-weighted fat saturated sequence or a STIR sequence to assess bone marrow edema.

Both flexor and extensor tenosynovitis will also be evaluated, according to the scoring system described by Haavardsholm et al. Tenosynovitis is scored at the level between the radioulnar joint and the hook of the hamate, thus including both wrist and finger tendons (60). Flexor and extensor tenosynovitis are evaluated semi-quantitatively in 10 different anatomical areas, graded from grade 0 to 3 (total score 0–30). The grading is based on the maximum width (in mm) of post-contrast enhancement within each anatomical area on axial T1-weighted MR images.

Joint space narrowing (JSN) will be scored according to the proposed OMERACT JSN scoring system for MRI (61). Each joint of the wrist and $1^{st}$ – $5^{th}$ MCP joint is scored 0-4, with 0 being no JSN and 4 being total ankylosis.

### *15.13 Appendix 13 WPAI (Norwegian translations available)*

**Work Productivity and Activity Impairment Questionnaire: Rheumatoid arthritis V2.0 (WPAI:RA)**

The following questions ask about the effect of your rheumatoid arthritis on your ability to work and perform normal daily activities. *Please fill in the blanks or circle a number, as indicated.*

1. Are you currently in paid employment? _____ NO ___ YES
   *If 'NO', tick "NO" and skip to question 6.*

The next questions refer to the **past seven days**, not including today.

2. During the past seven days, how many hours did you miss from work because of problems associated with your rheumatoid arthritis? *Include hours you missed on sick days, times you went in late, left early, etc., because of your rheumatoid arthritis. Do not include time you missed to participate in this study.*

   _____ HOURS

3. During the past seven days, how many hours did you miss from work because of any other reason, such as annual leave, holidays, time off to participate in this study?

   _____ HOURS

4. During the past seven days, how many hours did you actually work?

   _____ HOURS *(If "0", skip to question 6.)*

5. During the past seven days, how much did your rheumatoid arthritis affect your productivity while you were working?

*Think about days you were limited in the amount or kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual. If rheumatoid arthritis affected your work only a little, choose a low number. Choose a high number if rheumatoid arthritis affected your work a great deal.*

Consider only how much rheumatoid arthritis affected productivity while you were working.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Rheumatoid arthritis had no effect on my work | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Rheumatoid arthritis completely prevented me from working |

CIRCLE A NUMBER

6. During the past seven days, how much did your rheumatoid arthritis problems affect your ability to perform your normal daily activities, excluding your job?

*By normal activities, we mean the usual activities you perform, such as working around the house, shopping, childcare, exercising, studying, etc. Think about times you were limited in the amount or kind of activities you could perform and times you accomplished less than you would like. If rheumatoid arthritis affected your activities only a little, choose a low number. Choose a high number if rheumatoid arthritis affected your activities a great deal.*

Consider only how much rheumatoid arthritis affected your ability to perform your normal daily activities, excluding your job.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Rheumatoid arthritis had no effect on my daily activities | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Rheumatoid arthritis completely prevented me from performing my daily activities |

CIRCLE A NUMBER

WPAI:RA V2.0 (English for the UK)

## 15.14 Appendix 14 ARCTIC Treatment regimen

| Visit (months) | Treatment if no response (if response continue treatment at present step, see fig 2) |
|---|---|
| 1 (0) | **A. Monotherapy* + Prednisolone:**<br>1. Methotrexate 15 mg/week, increase by 2,5 mg every 2nd week to target dose 20 mg/week, i.e. week 1+2 15mg, week 3+4 17.5 mg, week 5-8 20 mg (optional reduced dosage starting scheme for patients at risk for side effects: week 1 10 mg, week 2 12,5mg, week 3 15 mg, week 4 17,5mg, week 5-8 20 mg)<br>2. Concomitant folic acid 5 mg/week (1mg 5/7 days or 5 mg x 1/week)<br>3. Prednisolone 15 mg week 1, 10 mg week 2, 7,5 mg week 3, 5 mg week 4+5, 2,5 mg week 6+7<br>4. Calcium supplement 1000mg x 1 (while on prednisolone) |
| 2 (1) | **A. Monitor start-up regimen (no changes in medication allowed unless due to AE)***<br>Joint injections allowed as indicated according to treatment arm. |
| 3 (2) | **A. Optimize monotherapy***<br>Increase Methotrexate to 25-30 mg/week<br>*Or increase SSZ/HCL/leflunomide dose* |
| 4 (3) | **A. Monitor start-up regimen (no changes in medication allowed unless due to AE)***<br>Joint injections allowed as indicated according to treatment arm. |
| 5 (4) | **B.** Triple combination therapy (or other combination therapy if MTX not tolerated):**<br>1. Add salazopyrine, step up over 4 weeks to 500mg 2 x 2 and<br>2. Add hydroxychlorochine 200mg 1 x 2 |
| 6 (6) | **B. Optimize triple combination therapy:****<br>Add Prednisolone 7,5 mg 1 x 1 |
| 7 (8) | **C. DMARD*** and 1st biologic**** (according to LIS guidelines):**<br>1. Highest tolerable dose MTX * and<br>2. Add 1st biologic<br>**Or SSZ/HCL/leflunomide if MTX not tolerated* |
| 8 (10) | **C. DMARD and 1st biologic:**<br>Adjust dose/interval of 1st biologic |
| 9 (12) | **D. DMARD*** and 2nd biologic (according to LIS guidelines):**<br>Switch to 2nd biologic<br>*(according to current LIS-guidelines)* |
| 10 (14) | **D. DMARD*** and 2nd biologic:**<br>Adjust dose/interval of 2nd biologic |
| 11 (16) | **E. DMARD*** and 3rd biologic (according to LIS guidelines):**<br>Switch to 3rd biologic<br>*(according to current LIS-guidelines)* |
| 12 (20) | **E. Optimize DMARD and 3rd biologic plus prednisolon:**<br>Adjust dose/interval of 3rd biologic and/or add prednisolon 7,5mg |
| 13 (24) | **F.** Continue medication according to standard clinical care |

** If MTX is not tolerated, switch to subcutaneous methotrexate (metoject), then continue according to scheme. In case of AE or not tolerated even in low dose subcutaneous, switch to salazopyrine or hydroxychlorochine monotherapy (standard dosage) if low disease activity, or leflunomide 20 mg in case of moderate or high disease activity (loading dose 40mg x 1 for 3 days, then 20 mg per day).*

*** In patients with high disease activity and risk factors for progressive joint destruction ( ACPA or RF-positive and either erosions on CR or baseline RAMRIS bone marrow edema score >2) a rescue option is available which includes moving to the next step, i.e. introduce 1st biologic (treatment C at visit #5, without prescribing treatment B).*

**** In case of no tolerance for any conventional DMARD, this can be omitted if the biologic drug chosen has indication for monotherapy (e.g. tociluzimab)*

***** Requirement for adding biologic: There must be objective signs of ongoing inflammation, i.e. either elevated ESR/CRP (>UNL, and not due to other disease/infection) or SJC>1 (or PD score >1 in US arm)*

### *15.15 Appendix 15 ARCTIC definition of decision rules*

These decision rules are to be applied at all visits except #2 and #4.

| | *Current DAS* | *No response** | *Response** | *Reached target (DAS < 1.6)* |
|---|---|---|---|---|
| **Conventional tight control** | ≤ 2.4 | Change of DAS < 0.6 | Change of DAS ≥ 0.6 | DAS < 1.6 and no swollen joints |
| | > 2.4 | Change of DAS < 1.2 | Change of DAS ≥ 1.2 | |
| **US tight control** | ≤ 2.4 | Change of DAS < 0.6 or <10% decrease of US total score | Change of DAS ≥ 0.6 and ≥ 10% decrease of US total score | DAS < 1.6 and no swollen joints and no joints with power Doppler synovitis |
| | > 2.4 | Change of DAS < 1.2 or <20% decrease of US total score | Change of DAS ≥ 1.2 and ≥ 20% decrease of US total score | |
| **Action** | | Change therapy | Continue current medication | Continue current medication** |

** Both in cases of response and no response should clinically swollen joints be i.a. injected with steroids when indicated, up to the maximum allowed dosage per visit (80 mg triamcinolone hexacetonid) . In group B joints with PD-signal on US is an additional target.*
*** If sustained remission ≥ 12 months, step-down to monotherapy MTX. If continued sustained response after this, decrease MTX by 2,5mg/week per 2 months.*

### *15.16 Flare questionnaire*

English version of the OMERACT flare questionnaire. A Norwegian translation will be available.

**OMERACT PRELIMINARY QUESTIONS for ASSESSMENT OF RA FLARE**
**VERSION 2.1**

**Section 1:**

1. Since your last visit is your rheumatoid arthritis (RA): please check box

| much worse – | worse – | slightly worse – | the same – | a little better – | better – | much better |
|---|---|---|---|---|---|---|
| ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

2. Are you having a Flare of your rheumatoid arthritis at this time?

Yes ☐ No ☐
If no, Please go to section 3

3. **If yes**, please circle a number to judge the severity of your flare during the last week.

No Flare 0 1 2 3 4 5 6 7 8 9 10 Extremely Bad

4. **If yes,** how long has this flare of your RA been going on?

☐ 1-3 days ☐ 4-7 days ☐ 8-14 days ☐ >14days

**Section 2. As a result of this "Flare", what did you do? (Tick/Check as many boxes as apply):**

☐ I didn't do anything different

☐ I reduced the amount of activities I did or/and rested more

☐ I avoided doing activities I had planned to do

☐ I tried to manage my flare myself, without medications (e.g. massage, warm/cold packs, exercise, had an extra visit to my physiotherapist, etc)

☐ I took more painkillers (analgesics) such as paracetamol, acetaminophen, codeine, tramadol, hydrocodone, oxycodone, etc or extra anti-inflammatory medications (NSAIDS) such as ibuprofen, ketoprofen, naproxen, celecoxib, meloxicam, etc

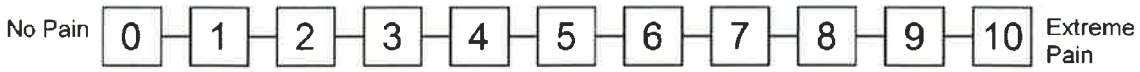
☐ I took more steroid tablets such as prednisone, methylprednisolone

☐ I asked for help from my Rheumatology Nurse, primary care provider, or my Rheumatologist

**Section 3: Describing your Rheumatoid Arthritis:**

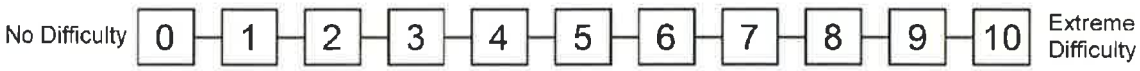
1. **PAIN:**

Circle the number that best describes the pain you have felt due to your rheumatoid arthritis during the last week:




**2. FUNCTION:**

Circle the number that best describes the difficulty you have had in doing physical activities (such as using your hands, walking or running, dressing, preparing meals etc) due to your rheumatoid arthritis during the last week:

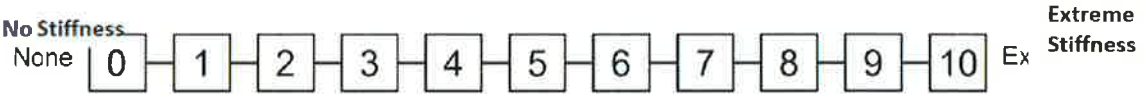



**3. FATIGUE:**

Circle the number that best describes how much fatigue you have felt due to your rheumatoid arthritis during the last week:

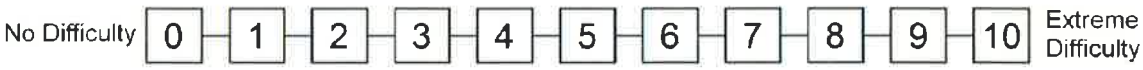

**4. STIFFNESS:**

Circle the number that best describes the stiffness (all over or/and in your joints) you have felt due to your rheumatoid arthritis during the last week:



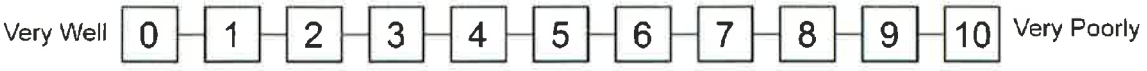

**5. PARTICIPATION:**

Considering how active your rheumatoid arthritis has been, please circle the number that best describes how much difficulty you have had during the last week, taking part in activities such as work, family life, social events that are normal for you:




**6. COPING/MANAGING YOUR LIFE:**

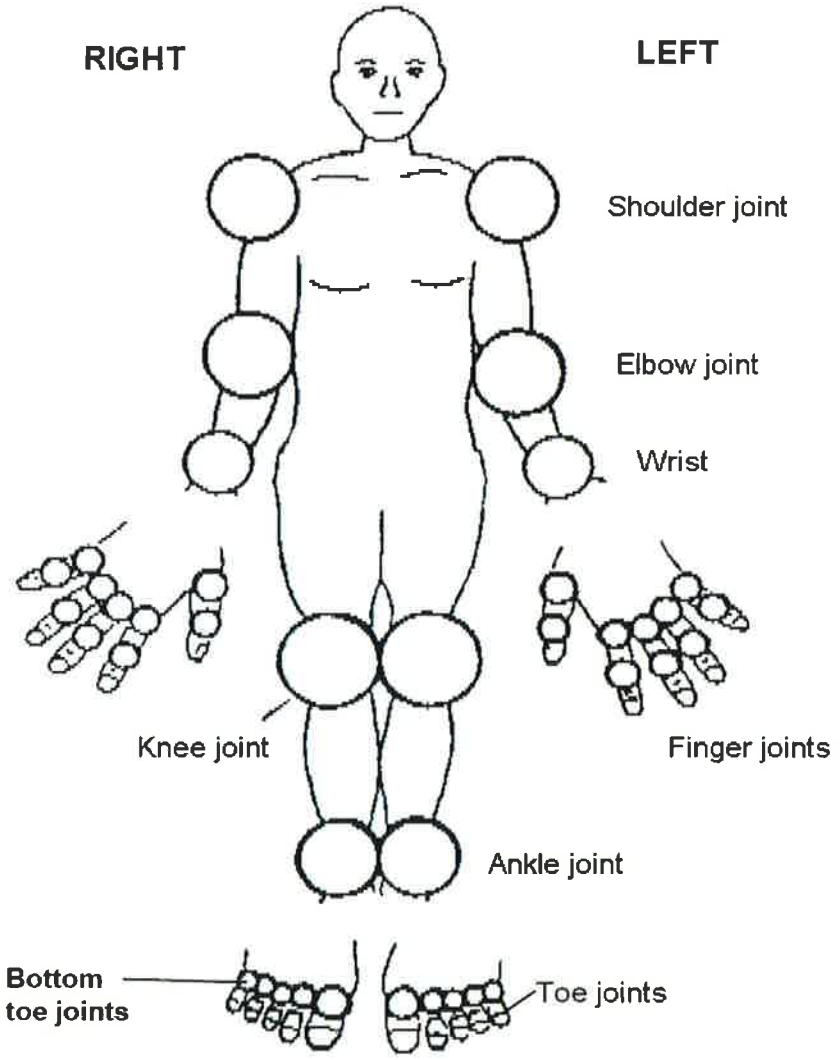
Considering your rheumatoid arthritis overall, please circle how well you have coped (managed, dealt with, made do) with your disease during the last week:



**Section 4: Please now evaluate your joints and mark on the joint pictures, which ones are tender or painful, and which ones are swollen at this time.**

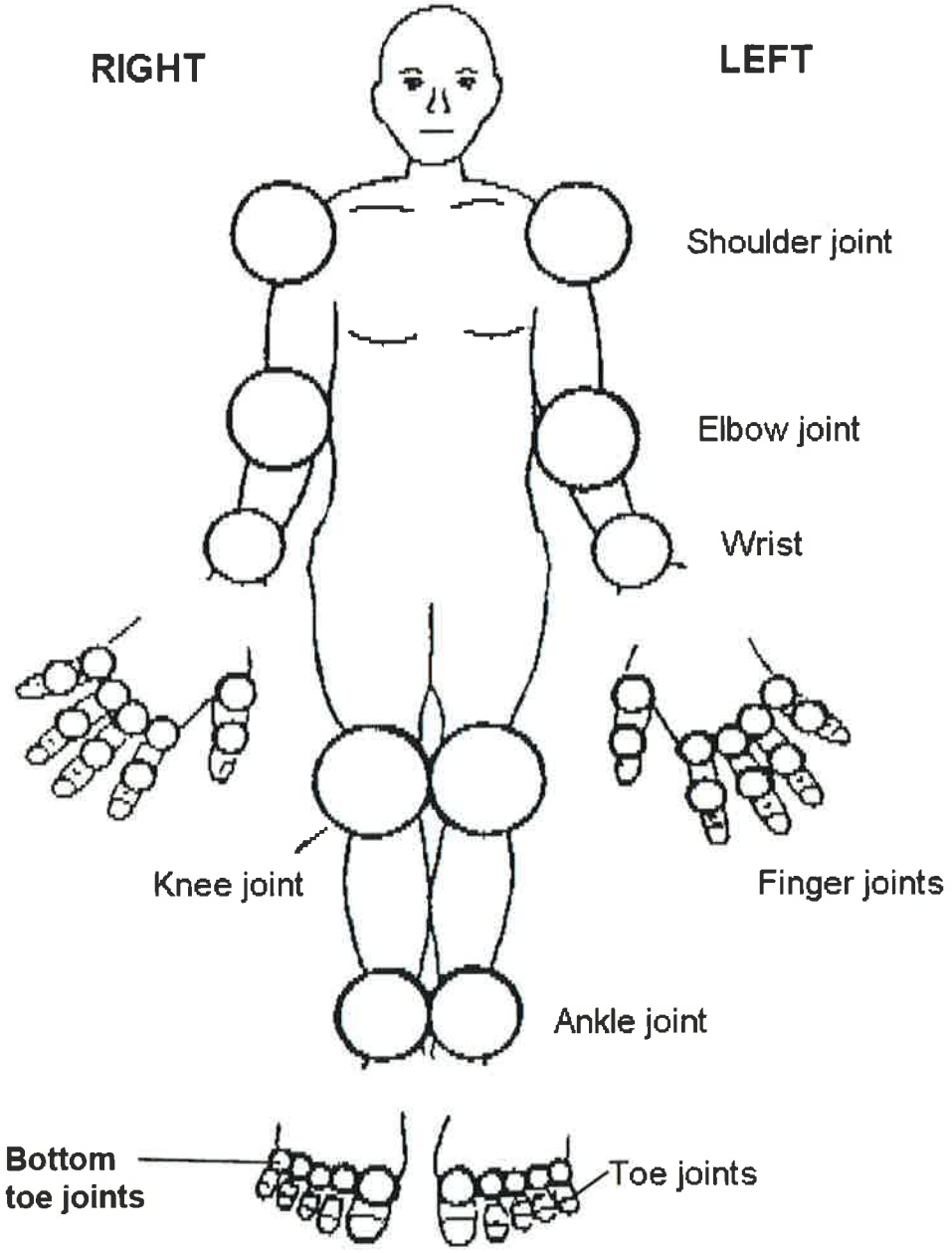
## Tender Joints

**Please indicate with a mark, on the picture below all the joints that are Tender at the present time:**



## Swollen Joints

**Please indicate with a mark, on the picture below all the joints that are <u>SWOLLEN</u> at the present time:**



## 16 Norwegian summary

# ARCTIC REWIND – norsk protokollsammendrag

| | |
|---|---|
| **Norsk kort tittel** | Behandling av RA i remision – **ARCTIC REWIND**-studien |
| **Deltakende sentre** | Revmatologisk avdeling ved:<br>Diakonhjemmet Sykehus AS (koordinerende senter/sponsor)<br>Universitetssykehuset i Nord-Norge<br>Haugesund Sanitetsforenings Revmatismesykehus AS<br>Helse Møre og Romsdal HF, Ålesund sjukehus<br>Haukeland Universitets-sykehus Bergen<br>Sørlandet Sykehus Kristiansand HF<br>Martina Hansens Hospital<br>Sykehuset Østfold HF Moss<br>Lillehammer Revmatismesykehus<br>Vestre Viken HF Drammen Sykehus<br>Privat spesialistpraksis Bendvold/Dovland Kristiansand<br>Revma Vestfold<br>Helgelandssykehuset Mo i Rana |
| **Hovedutprøver** | Prof. Espen A. Haavardsholm (MD, PhD), Diakonhjemmet Sykehus |
| **Prosjektleder** | Siri Lillegraven (MD, MPH, PhD), Diakonhjemmet Sykehus |
| **Klinisk ansvarlig** | Avd.sjef professor dr. med. Tore K. Kvien, Diakonhjemmet Sykehus |
| **Rådgivere, metodologi** | Professor Désirée van der Heijde, MD PhD, Professor Daniel H. Solomon, MD MPH, Professor Robert Landewé, MD PhD, Professor Ivar S. Kristiansen, MD PhD MPH |
| **Statistiker** | Inge C. Olsen, PhD, Joseph Sexton PhD |
| **Monitor (GCP)** | Avdeling for klinisk forskningsstøtte, OUS (Helse Sør-Øst) og Innovest AS (Haukeland), Klinisk forskningsavdeling, UNN (Tromsø), Avdeling for forskning og utvikling, Helse Møre og Romsdal (Ålesund) |
| **Fase** | IV |
| **Populasjon** | Pasienter med revmatoid artritt i henhold til ACR/EULAR 2010 kriterier, og vedvarende remisjon i ett år ELLER inklusjon i ARCTIC-studien |
| **Målsetting** | Å undersøke effekten av nedtrapping og seponering av sykdomsmodifiserende legemidler på sykdomsaktivitet hos leddgiktspasienter som er i vedvarende klinisk remisjon |
| **Design** | Randomisert, åpen, kontrollert, parallell-gruppe, multisenter fase 4, non-inferiority strategistudie. |
| **Antall pasienter** | 360 |
| **Tidsplan** | Inklusjonsperiode: 01.03.13 til 30.06.18, deretter 36 måneders oppfølgning. |
| **Behandlingsregime** | Biologiske og syntetiske DMARDs i varierende doseringer og kombinasjoner. |
| **Effektmål** | Sykdomsoppbluss, ACR kjernesett av utfallsmål, bildediagnostiske undersøkelser og selvrapporterte utfallsmål |
| **Bivirkningsregistrering** | Puls/blodtrykk, hematologi, klinisk kjemi, rapportering av bivirkninger og alvorlige bivirkninger etter ICH Guidelines (med noen modifikasjoner). |

## Summary of changes in the protocol from the first to the final version

**Version 1_2** – Initial protocol submitted to the Regional Ethics Committee (Dec 12 2012) and to the Norwegian Medicines Agency (Dec 13 2012)

**Version 2_1** – Updated protocol submitted to the Regional Ethics Committee (Feb 27 2013) and to the Norwegian Medicines Agency (Feb 27 2013). Main changes made: Update to signature page, update in description of how suspected unexpected serious adverse reactions (SUSARs) should be reported, specification that the non-inferiority analyses will be done in the PP population and the superiority analyses in the ITT population as appropriate.

**Version 3_1** – Updated version number and date on March 25 2013, otherwise unchanged from previous version.

**Version 4_0** – Submitted to the Norwegian Medicines Agency on April 23 2017 and to the Regional Ethics Committee on May 04 2017. Updated version number, updates made to project group and eligible study centers, update of inclusion criteria in section 3.3 (inclusion criteria d, length of disease duration), clarification that the two randomized clinical trials will be analyzed separately (in section 7.1.

**Version 4_1** – Final version. Submitted to the Regional Ethics Committee on December 01 2017. Changes made: MRI at 36 months will be performed without gadolinium-based contrast agent.

# ARCTIC REWIND
# REmission in rheumatoid arthritis – assessing WIthdrawal of disease-modifying antirheumatic drugs in a Non-inferiority Design

## Statistical Analysis Plan Final Version 1.1, 18.06.2019
Analyses of patients who receive synthetic disease modifying antirheumatic drugs

**Protocol Identification Number:** DIA2012-1 / ver 4_1
**EudraCT Number:** 2012-005275-14

**APPROVAL PAGE**

**Biostatistician Diakonhjemmet Hospital**

Date 18/6 - 2019

Signature

*Joseph Sexton, PhD*

**Principal investigator**

Date 18/6-19

Signature

*Espen A. Haavardsholm, MD PhD*

**Project leader**

Date 18/6 - 2019

Signature

*Siri Lillegraven, MD MPH PhD*

# Abbreviations

| | |
|---|---|
| ARD | Adjusted Risk Difference |
| ARR | Adjusted Relative Risk |
| ATC | Anatomical Therapeutic Chemical |
| AUC | Area Under the Curve |
| BMD | Bone Mineral Density |
| BME | Bone Marrow Edema |
| BMI | Body Mass Index |
| CDAI | Clinical Disease Activity Index |
| CI | Confidence Interval |
| CRP | C-Reactive Protein |
| CSR | Clinical Study Report |
| DAS | Disease Activity Score |
| DMARD | Disease Modifying Anti-Rheumatic Drug |
| eCRF | Electronic Case Report Form |
| EOT | End of Treatment |
| ESR | Erythrocyte Sedimentation Rate |
| EULAR | European League Against Rheumatism |
| FAS | Full Analysis Set |
| FDA | U.S. Food and Drug Administration |
| GS | Grey Scale |
| HAQ | Health Assessment Questionnaire |
| HR | Hazard Ratio |
| HRQoL | Health related quality of life |
| IDMC | Independent Data Monitoring Committee |
| ITT | Intention to Treat |
| JSN | Joint Space Narrowing |
| KM | Kaplan Meier |
| MCP | Metacarpophalangeal |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MRI | Magnetic resonance image |
| MTP | Metatarsophalangeal |
| MTX | Methotrexate |
| NRS | Numeric rating scale |
| PD | Power Doppler |
| PGA | Patient's Global Assessment of disease activity |
| PH | Proportional Hazards |
| PhGA | Physician's Global Assessment of disease activity |
| PIP | Proximal interphalangeal |
| PP | Per Protocol |
| PT | Preferred Term |
| QoL | Quality of Life |

| | |
|---|---|
| RA | Rheumatoid Arthritis |
| RAI | Ritchie Articular Index |
| RAID | The Rheumatoid Index of Disease |
| SAE | Serious Adverse Event |
| SAP | Statistical Analysis Plan |
| SC | Sternaclavicular |
| SD | Standard Deviation |
| SDAI | Simplified Disease Activity Index |
| SE | Standard Error |
| SJC | Swollen Joint Count |
| SOC | System Organ Class |
| TEAE | Treatment Emerging Adverse Events |
| TFL | Tables Figures Listings |
| TJC | Tender Joint Count |
| US | Ultrasound |
| VAS | Visual Analogue Scale |
| vdHSS | Sharp van derHeijde Score |
| WHO | World Health Organisation |
| WPAI | Work Productivity and Activity Impairment Questionnaire |

# TABLE OF CONTENTS

# 1. INTRODUCTION

This document describes the planned data summaries and statistical analyses to be performed for the primary analyses/first study period (0-12 months) in patients receiving synthetic disease modifying antirheumatic drugs (DMARDs) in the ARCTIC REWIND trial (REmission in rheumatoid arthritis – assessing WIthdrawal of disease-modifying antirheumatic drugs in a Non-inferiority Design, Clinical Trial Protocol DIA2012-1). It is intended to supplement the study protocol, which contains details regarding the objectives and design of the study.

## 1.1 Study Objectives

### 1.1.1 Primary Objective

The primary objective of this treatment strategy study is to assess the effect of tapering and withdrawal of DMARDs on disease activity in RA patients in sustained remission.

### 1.1.2 Secondary Objectives

The secondary objectives of the complete ARCTIC REWIND study are listed below, the analysis of secondary objectives will not be described in detail in the statistical analysis plan (SAP).

a) To study predictors of successful DMARD reduction and discontinuation in RA remission
b) To evaluate the cost-effectiveness of alternative treatment strategies in stable RA remission
c) To assess whether the choice of DMARD strategy in RA remission influences the level of magnetic resonance imaging (MRI) and ultrasonography detected inflammation
d) To assess joint damage in patients in sustained RA remission who continue to receive stable DMARD treatment and patients who receive reduced DMARD treatment
e) To provide long-term follow-up data on patients included in the original ARCTIC trial
f) To assess relationships between treatment, inflammation, physical function and joint damage in RA remission
g) To examine adverse event rates in different DMARD strategies in RA remission
h) To examine the value of imaging information in treatment decision making in sustained RA remission
i) To study how serum drug levels and/or anti-drug antibodies are associated to drug efficacy in RA remission
j) To assess the performance of definitions of RA flare
k) To assess differences in success of DMARD reduction and discontinuation in patients according to disease duration
l) To assess outcomes after flares, including response to increased treatment

## 1.2 Study Design

ARCTIC REWIND is a 36-month, randomized, open-label, phase IV, parallel-group, multi-center, non-inferiority study to evaluate the effect of DMARD dose reduction in RA subjects who have achieved sustained remission.

ARCTIC REWIND includes two separate randomized clinical trials, one assessing tapering and withdrawal of synthetic DMARDs (sDMARDs), and one assessing tapering and withdrawal of tumor necrosis factor inhibitors (TNFi). The current SAP describes the planned analyses of the primary endpoint and main secondary endpoints in the sDMARD study (period 1, 0-12 months). The SAP will later be updated with the analyses for the second (12-24 months) and extended follow-up study period

(24-36 months). A separate SAP will outline the analyses of the TNFi study. The design of the sDMARD trial is illustrated in **Figure 1**.

Patients included in ARCTIC REWIND are followed for 36 months, with visits every 4 months. The study is divided into three periods as outlined in **Table 1.**

### 1.2.1 Dosage and drug administration

Patients in the synthetic DMARD ARCTIC REWIND study will at inclusion either receive a single synthetic DMARD in monotherapy, or multiple synthetic DMARDs in combination therapy. Patients cannot have received any biologic DMARDs during the previous 12 months. At baseline patients will be randomized to either continued stable synthetic DMARD treatment (S1) or half dose synthetic DMARD treatment (S½). Patients on combinations of synthetic DMARDs who are randomized to the S½ arm will continue combination therapy, but each DMARD will be reduced to half dose of the baseline dose. If the baseline dose could not be changed to exactly half dose (e.g. methotrexate 17.5 mg/week), it was left to the clinician's preference if the lowest or highest potential dose was chosen (e.g., for the example above methotrexate 7.5 mg/week or methotrexate 10 mg/week). Examples of the half dose regimen for each synthetic DMARD is outlined in **Table 2**: Dosing regimens in ARCTIC REWIND.

In Period 2, patients in arm S1 who have not had a flare will continue their treatment unchanged. Patients in arm S½ who have not experienced a flare will be randomized 1:1 to either continue the S½ treatment or discontinue their synthetic DMARD(s) (S0). These patients will at this second randomization have been in sustained remission for at least two years.

Patients randomized to S1 at baseline will continue this treatment regimen throughout the study as long as the treatment result is satisfactory. If the patient in the S1 arm experiences a flare, treatment can be escalated according to the physician's preference, preferentially as outlined in the ARCTIC treatment guidelines (Appendix 16.14 in the protocol).

If a patient in treatment arms S½ or S0 experiences a flare, he or she will return to the full dose of the study medication (the treatment the patient was receiving at baseline). If necessary, treatment could be further escalated according to current recommendations and the physician's preference.

**Figure 1:** Illustration of ARCTIC REWIND sDMARD study design.

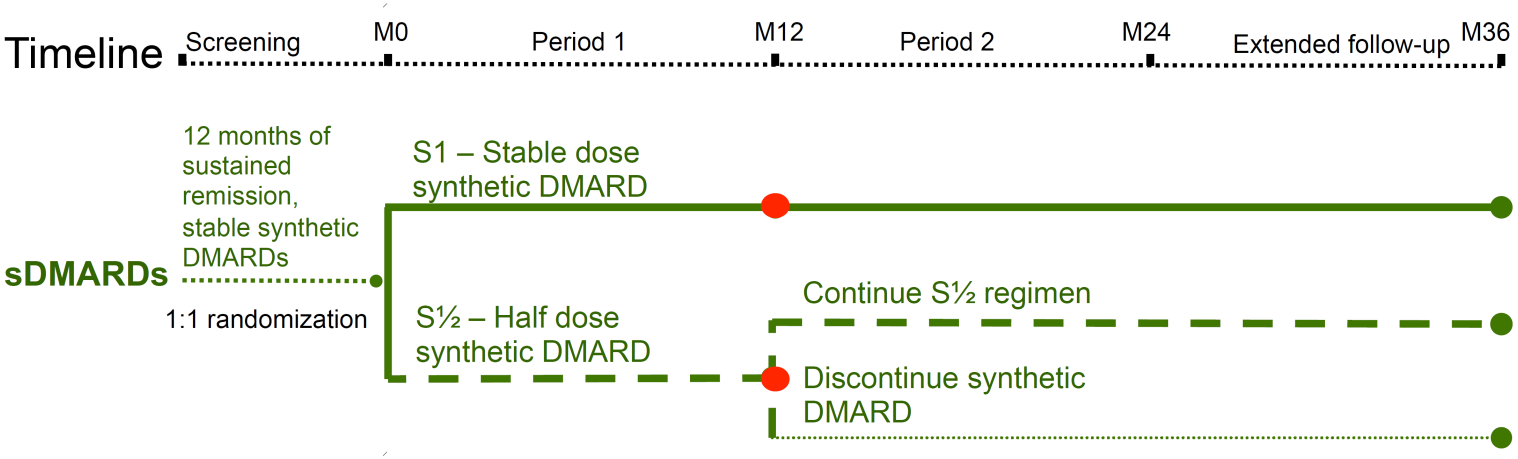


**Table 1:** Overview of ARCTIC REWIND sDMARD treatment periods

| ***Period 1***<br>*Baseline (Visit 1) to 12 months (Visit 4)* | ***Period 2***<br>*12 months (Visit 4) to 24 months (Visit 7)* | ***Period 3***<br>*24 months (Visit 7) to 36 months (Visit 10)* |
|---|---|---|
| Patients in will be randomized 1:1 into continued stable synthetic DMARD(s) (S1) or half dose synthetic DMARD(s) (S½). The reduction in synthetic DMARD dose will be done according to the plan in **Table2.** | Patients in sustained remission on **S1** regimen (stable dosage) will continue this treatment. Patients in sustained remission on S½ treatment will be randomized into continued half dose synthetic DMARD treatment (S½→½) or discontinued synthetic DMARD treatment (S½→0). | Patients in sustained remission will continue to receive the treatment they were assigned to in Period 2. Patients are followed with regards to efficacy and safety measures. |

**Table 2:** Example of half dose regimens in ARCTIC REWIND

| ***Drug*** | ***Producer*** | ***Package*** | ***ATC*** | ***Standard full dosage*** | | ***1/2 dosage regimen*** | |
|---|---|---|---|---|---|---|---|
| | | | | **Dosage** | **Frequency** | **Dosage** | **Frequency** |
| **Synthetic DMARDs** | | | | | | | |
| **Methotrexate tablets** *(methotrexate)* | Pfizer | 2,5 mg no.100 | L04A X03 | 25mg p.o. | Weekly | 12,5mg p.o. | Weekly |
| **Metex prefilled syringe** *(methotrexate)* | Medac | 50mg/ml 0,5ml x 6 syringes | L01B A01 | 25mg s.c. | Weekly | | |
| **Metex prefilled syringe** *(methotrexate)* | Medac | 50mg/ml 0,25ml x 6 syringes | L01B A01 | | | 12,5mg s.c. | Weekly |
| **Salazopyrin EN tablets** *(sulfasalazine)* | Pfizer | 500mg No. 100 | A07E C01 | 1000mg x 2 p.o. | Daily | 500mg x 2 p.o. | Daily |
| **Plaquenil** *(hydroxychlorochine)* | Sanofi-Aventis | 200mg No. 100 | P01B A02 | 200mg x 2 p.o. | Daily | 200mg x 1 p.o. | Daily |
| **Arava** *(leflunomide)* | Sanofi-Aventis | 20mg No 100 | L04A A13 | 20mg p.o. | Daily | | |
| **Arava** *(leflunomide)* | Sanofi-Aventis | 10mg No 100 | L04A A13 | | | 10mg p.o. | Daily |

# 2. HYPOTHESES AND DECISION RULES

## 2.1 Statistical Hypothesis

For the rest of the statistical analysis plan, the term "study" is used to refer to the first 12 months of the sDMARD ARCTIC REWIND trial (period 1).

This study is designed to assess the non-inferiority of tapered sDMARD therapy compared to stable sDMARD therapy with regard to disease flares during 12 months of follow-up in patients who have been in stable remission for at least 12 months at baseline.

Here, a patient that does not experience a disease flare in the course of the study is referred to as a **non-failure.** The **treatment difference (Δ)** is defined to be the probability that a patient on stable sDMARD treatment (S1) is a non-failure minus the probability that a patient on tapered sDMARD therapy (S1/2) is a non-failure.

The null-hypothesis ($H_0$) is that tapered sDMARD therapy is inferior to stable sDMARD therapy, here defined as the treatment difference being equal to or in excess of 20 percentage points. The alternative hypothesis ($H_1$) is that the treatment difference is less than 20 percentage points. Formally stated as:

$H_0$: $\Delta \geq 20\%$
versus
$H_1$: $\Delta < 20\%$

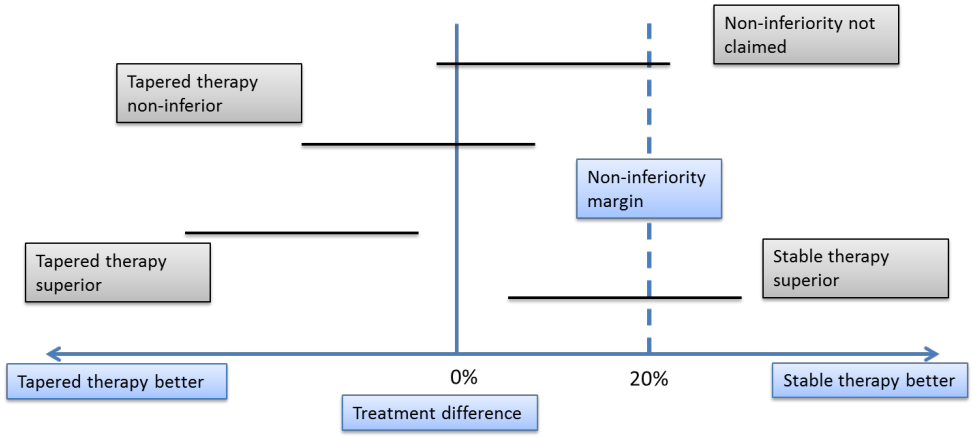
If the primary null hypothesis is rejected, we will assess the superiority of tapered therapy relative to stable therapy. If it is not rejected, the superiority of stable over tapered therapy will be assessed.

## 2.2 Statistical Decision Rule

The hypothesis test will be carried out at the 0.025 (alpha) level. Operationally, letting ($\Delta_L$, $\Delta_U$) denote the 95% confidence interval for Δ, we will conclude that tapered sDMARD therapy is non-inferior to stable sDMARD therapy if $\Delta_U$<20%. The confidence interval will equal the acceptance region for a two-sided hypothesis test.

Furthermore, if the lower end point of this confidence interval exceeds zero ($\Delta_L$ >0) we will conclude that stable sDMARD therapy is superior to tapered therapy, and conversely if the upper end point is less than zero ($\Delta_U$ <0) we will conclude that tapered therapy is superior.

**Figure 2.1 Illustration of statistical decision rule (bars indicate 95% confidence limits)**



# 3. ANALYSIS SETS

## 3.1 Enrolled

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

## 3.2 Full Analysis Set

The Full Analysis Set (FAS) will be defined as all patients randomly assigned to a treatment group, and have started the allocated intervention defined as having completed at least one regular visit after the baseline visit.

## 3.3 Safety Analysis Set

The Safety Analysis Set will include all patients who completed at least one regular visit after the baseline visit.

## 3.4 Per Protocol Analysis Set

The Per Protocol Analysis Set (PPS) will include all randomised patients meeting the study entry criteria and with no major protocol deviations affecting the treatment efficacy (see SAP section 3.8).

## 3.5 Flare Analysis Set

The Flare Analysis Set will include any patient who was include in the full analysis set and experiences a flare during the study

## 3.6 Tapering Analysis Set

The tapering Analysis Set (TAS) will include any patient who was in the per protocol analysis set and randomized to tapering therapy (S1/2).

## 3.7 Treatment Misallocation

If patients were:

- randomized but not treated: patient will appear on the patient evaluation table as randomized but not treated; this is the extent of how the patient will be reported.
- treated but not randomized: then by definition the patient will be excluded from both the efficacy and safety analyses since randomized treatment is missing
- randomized but did not follow protocol according to allocation: then they will be reported under their randomized treatment group for all efficacy and safety analyses as part of the FAS and safety analyses, but omitted from the PPS

Note that treated in this section is defined as having completed at least one post-baseline regular visit.

## 3.8 Protocol Deviation

The following sections describe any protocol deviations that relate to the statistical analyses and forms the requirement for exclusion from the PPS.

### 3.8.1 Deviations to inclusion and/or exclusion criteria

Not fulfilling inclusion and exclusion criteria will be considered a protocol deviation.

### 3.8.2 Deviations assessed Post-Randomization

Only protocol deviations thought to affect the efficacy of the assessed DMARD treatment will be considered in the SAP. Each of these cases will be reviewed by the team and a clinical judgment made in each particular circumstance as to whether efficacy would have been affected in the case of these specific classes of protocol deviations assessed post randomization:

- Patients who did not follow the treatment regimens outlined in section 1.2.1
- Patients who withdrew or was withdrawn during the study

# 4. DEFINITIONS AND DERIVED VARIABLES

In this section we outline the variables used in the study, including variables that will be used in subsequent analyses of secondary objectives not covered in the primary publication. For all clinically planned measures, visits should occur within a window of the scheduled visit, see table below for definitions.

| Visit Label | Target Day | Definition (Day window) |
|---|---|---|
| Screening | -1 | Prior to Day 0 |
| V1. Baseline | Day 0 (Randomization) | Day 0 |
| V2. Month 4 | 122 | Day 1 to 183 |
| V3. Month 8 | 243 | Day 184 to 304 |
| V4. Month 12 | 365 | Day 304 to 426 |
| V5. Month 16 | 487 | Day 426 to 548 |
| V6. Month 20 | 608 | Day 549 to 669 |
| V7. Month 24 | 730 | Day 670 to 791 |
| V8. Month 28 | 852 | Day 792 to 913 |
| V9. Month 32 | 972 | Day 914 to 1034 |
| V10. Month 36 | 1095 | Day 1035 to 1156 |

If two or more visits fall into the same window, keep the one closest to the Target Day. If two visits are equal distant from the Target Day in absolute value, the later visit should be used. If two visits are within the window of visit 4, the visit with inclusion of radiology and biobanking should be used.

## 4.1 Change from baseline

Change from baseline (Δ) = time-point value - baseline value.
% change from baseline (%Δ) = [(time-point value – baseline value) / baseline value] *100%

## 4.2 Inflammation parameters

Inflammation parameters include the Erythrocyte Sedimentation Rate (ESR) in mm/h and high sensitivity C-reactive protein (CRP) in mg/L. ESR is assessed by the Westergren method.

## 4.3 Synthetic Disease Modifying Anti-Rheumatic Drugs

Synthetic DMARDs include, for this study, the following drugs: Methotrexate, sulfasalazine, leflunomide, hydroxychloroquine.

## 4.4 Joint Counts

The Ritchie Articular Index (RAI) is a long-standing approach to a graded assessment of the tenderness of 26 joint regions, based on summation of joint responses after applying firm digital pressure. Four grades can be used: 0, patient reported no tenderness; +1, patient complained of pain; +2, patient complained of pain and winced; and +3, patient complained of pain, winced, and withdrew. Thus, the index ranges from 0 to 3 for individual measures and the sum 0 to 78 overall, with higher scores indicating more tenderness. Certain joints are treated as a single unit, such as the metacarpal-phalangeal and proximal interphalangeal joints of each hand and the metatarsophalangeal joints of each foot. For example, the maximum score for the five metacarpophalangeal joints of the right hand would be 3, not 15.

Swollen joint counts are performed on 44 joints, with total joint count ranging from 0 to 44. This is denoted the 44 swollen joint count (SJC44).

In addition, tender and swollen joint counts on 28 joints are denoted TJC28 and SJC28. A tender joint is defined as a joint with RAI > 0. For the MCP and IP/PIP regions we use the following definition to estimate TJC28: if RAI = 0, TJC(region) = 0; if RAI = 1, TJC(region) = 2; if RAI = 2, TJC(region) = 3: if RAI = 3, TJC(region) = 4.

See Table 4.1 Overview of joint counts for an overview of assessed joints and how they are scored for the different joint counts

**Table 4.1 Overview of joint counts**

| Joints | RAI left | RAI right | SJC44 left | SJC44 right | SJC28 left | SJC28 right |
|---|---|---|---|---|---|---|
| **Cervical spine (Neck)** | 0-3 | | NA | NA | NA | NA |
| **Temporomandibular (Jaws)** | 0-3 | | NA | NA | NA | NA |
| **Sternoclavicular (SC)** | 0-3 | | 0-1 | 0-1 | NA | NA |
| **Acromioclavicular (AC)** | 0-3 | | 0-1 | 0-1 | NA | NA |
| **Shoulder*** | 0-3 | 0-3 | 0-1 | 0-1 | 0-1 | 0-1 |
| **Elbow*** | 0-3 | 0-3 | 0-1 | 0-1 | 0-1 | 0-1 |
| **Wrist*** | 0-3 | 0-3 | 0-1 | 0-1 | 0-1 | 0-1 |
| **Metacarpophalangeal (MCP)*** | | | | | | |
| **- First (MCP1)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Second (MCP2)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Third (MCP3)** | 0-3 | 0-3 | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Fourth (MCP4)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Fifth (MCP5)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **Proximal interphalangeal (IP/PIP)*** | | | | | | |
| **- First (IP1)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Second (PIP2)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Third (PIP3)** | 0-3 | 0-3 | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Fourth (PIP4)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **- Fifth (PIP5)** | | | 0-1 | 0-1 | 0-1 | 0-1 |
| **Hip** | 0-3 | 0-3 | NA | NA | NA | NA |
| **Knee*** | 0-3 | 0-3 | 0-1 | 0-1 | 0-1 | 0-1 |
| **Ankle** | 0-3 | 0-3 | 0-1 | 0-1 | NA | NA |
| **Talocalcaneal** | 0-3 | 0-3 | NA | NA | NA | NA |
| **Tarsus** | 0-3 | 0-3 | NA | NA | NA | NA |
| **Metatarsophalangeal (MTP)** | | | | | | |
| **- First (MTP1)** | | | 0-1 | 0-1 | NA | NA |
| **- Second (MTP2)** | | | 0-1 | 0-1 | NA | NA |
| **- Third (MTP3)** | 0-3 | 0-3 | 0-1 | 0-1 | NA | NA |
| **- Fourth (MTP4)** | | | 0-1 | 0-1 | NA | NA |
| **- Fifth (MTP5)** | | | 0-1 | 0-1 | NA | NA |

NA: Not assessed; *: Included in the 28-joint Disease Activity Score (DAS28)

## 4.5 ACR core data set

The American College of Rheumatology (ACR) definition of response includes tender and swollen joint counts, visual analogue scales (VAS) for pain, patient and investigator global assessment of disease activity, patient-assessed disability by the Health Assessment Questionnaire (HAQ) and acute phase response (ESR or high sensitivity CRP).

## 4.6 ACR response

If a patient experiences a flare and treatment is escalated, the ACR response rates ACR20, ACR50, ACR70 and ACR90 as well as ACR remission rates will be calculated.

An ACR20 response is defined if the following criteria are fulfilled:

- 20% improvement in RAI
  *AND*
- 20% improvement in swollen joint count 44
  *AND*
- 20% improvement in at least 3 of 5 other core set items

The other core set items consist of:

- Investigator global assessment of disease activity
- Patient global assessment of disease activity
- Patient pain

- Disability
- ESR/CRP

ACR50, ACR70 and ACR90 are defined in a similar manner with 50%, 70% and 90% improvement, respectively. In the ARCTIC REWIND study, VAS will be used to assess pain and patient/investigator global assessment of disease activity, and PROMIS Physical Function raw score will be used to assess disability. High sensitivity CRP will be used as primary measure of inflammation, while ESR will be used if CRP is not available. All improvements will be % change from baseline.

Time to ACR20/50/70/90 response = Date of first visit with ACR20/50/70/90 response – date of flare +1

### 4.6.1 US Food and Drug Administration (FDA) major clinical response
Requires achieving ACR70 response at the current visit and at each visit within the previous 6 months.

## 4.7 Disease Activity
### 4.7.1 DAS
Disease Activity Score (DAS) includes the RAI, the 44 swollen joint counts, the ESR and the patient's global assessment of disease activity on a VAS 0-100 mm (PGA).

The DAS is calculated as follows:
**DAS = 0.54*sqrt(RAI) + 0.065*(SJC44) + 0.33*Ln(ESR) + 0.0072*PGA**

If values of ESR and/or PGA are missing, the following formulas are used
**DAS = 0.54*sqrt(RAI) + 0.065*(SJC44) + 0.33*Ln(ESR) + 0.22**
**DAS = 0.54*sqrt(RAI) + 0.065*(SJC44) + 0.17*Ln(CRP+1) + 0.0072*PGA + 0.45**
**DAS = 0.54*sqrt(RAI) + 0.065*(SJC44) + 0.17*Ln(CRP+1) + 0.65**

According to DAS, the following cut-points are used:
High disease activity: DAS > 3.7
Moderate disease activity: 3.7 ≥ DAS>2.4
Low disease activity: 2.4 ≥ DAS ≥ 1.6
In remission: DAS < 1.6

### 4.7.2 DAS28
The 28-joint Disease Activity Score (DAS28) includes TJC28, SJC28, ESR and PGA.

The DAS28 is calculated as follows:
**DAS28 = 0.56*sqrt(TJC28) + 0.28*sqrt(SJC28) + 0.70*Ln(ESR) + 0.014*PGA**

If values of ESR and/or PGA are missing, the following formulas are used:
**DAS28 = [0.56*sqrt(TJC28) + 0.28*sqrt(SJC28) + 0.70*Ln(ESR)]*1.08 + 0.016**
**DAS28 = 0.56*sqrt(TJC28) + 0.28*sqrt(SJC28) + 0.36*Ln(CRP+1) + 0.014*PGA + 0.96**
**DAS28 = [0.56*sqrt(TJC288) + 0.28*sqrt(SJC28) + 0.36*Ln(CRP+1)]*1.10 + 1.15**

According to DAS28, the following cut-points are used:
High disease activity: DAS28 > 5.1
Moderate disease activity: 5.1 ≥ DAS28>3.2
Low disease activity: 3.2 ≥ DAS28 ≥ 2.6
In remission: DAS28 < 2.6

### 4.7.3 CDAI
The Clinical Disease Activity Index (CDAI) includes TCJ28, SJC28, PGA in addition to the treating physician's global assessment of disease activity on a VAS 0-100 mm (PhGA).

The CDAI is calculated as follows:
**CDAI=TCJ28 + SJC28 + PGA/10 + PhGA/10**

According to CDAI, the following cut-points are used:
High disease activity: CDAI > 22.0
Moderate disease activity: 22.0 ≥ CDAI>10.0
Low disease activity: 10.0 ≥ CDAI > 2.8
In remission: CDAI ≤ 2.8

### 4.7.4 SDAI

The Simplified Disease Activity Index (SDAI) includes TCJ28, SJC28, PGA, PhGA and CRP.

The SDAI is calculated as follows:
**SDAI=TCJ28 + SJC28 + PGA/10 + PhGA/10 + CRP/10**

According to SDAI, the following cut-points are used:
High disease activity: SDAI> 26.0
Moderate disease activity: 26.0 ≥ SDAI>11.0
Low disease activity: 11.0 ≥ SDAI > 3.3
In remission: SDAI ≤ 3.3

## 4.8 EULAR response

If a patient has experienced a flare, and treatment has been escalated, EULAR response will be calculated. The European League Against Rheumatism (EULAR) response rates will be calculated. A EULAR response is defined by the state and change in DAS and DAS28, and categorized into good, moderate and none using the following definitions:

**Table 4.2** EULAR DAS response

| | Change from relevant visit (e.g. flare visit) in DAS | | |
|---|---|---|---|
| DAS at time-point | ΔDAS ≤ - 1.2 | -1.2 < DAS < -0.6 | DAS ≥ 0.6 |
| DAS ≤ 2.4 | Good | Moderate | None |
| 2.4 < DAS ≤ 3.7 | Moderate | Moderate | None |
| DAS > 3.7 | Moderate | None | None |

**Table 4.3** EULAR DAS28 response

| | Change from relevant visit (e.g. flare visit) in DAS28 | | |
|---|---|---|---|
| DAS28 at time-point | ΔDAS28 ≤ - 1.2 | -1.2 < DAS28 < -0.6 | DAS28 ≥ 0.6 |
| DAS28 ≤ 3.2 | Good | Moderate | None |
| 3.2 < DAS28 ≤ 5.1 | Moderate | Moderate | None |
| DAS28 > 5.1 | Moderate | None | Noen |

Time to EULAR response = Date of first visit with EULAR response – date of flare +1

## 4.9 The van der Heijde modified Sharp Score

The van der Heijde modified Sharp Score (vdHSS) is a score of erosion and joint space narrowing (JSN) based on radiographs of hands and feet. The score for JSN ranges from 0 to 120 in the hands and from 0 to 48 in the feet. The total vdHSS score is the sum of scores for erosions and JSN, the maximum score being 448.

Radiographic progression is defined as a ΔvdHSS of ≥ 1 unit/year. I.e. a radiographic progression after one year (visit 4) is defined as a ΔvdHSS of ≥ 1 unit/year. Rate of progression is defined as ΔvdHSS divided by time (years). Progression rates will in addition be dichotomized according to ≥ 0.5, ≥ 2.0 and ≥ 5.0 units per year.

**Table 4.4 Overview of the van der Heijde modified Sharp Score**

| Area | Joints | Erosion left | Erosion right | JSN left | JSN right |
|---|---|---|---|---|---|
| **Hand** | Metacarpophalangeal (MCP) | | | | |
| | - First (MCP1) | 0-5 | 0-5 | 0-4 | 0-4 |
| | - Second (MCP2) | 0-5 | 0-5 | 0-4 | 0-4 |
| | - Third (MCP3) | 0-5 | 0-5 | 0-4 | 0-4 |
| | - Fourth (MCP4) | 0-5 | 0-5 | 0-4 | 0-4 |
| | - Fifth (MCP5) | 0-5 | 0-5 | 0-4 | 0-4 |
| | Proximal interphalangeal (IP/PIP) | | | | |
| | - First (IP1) | 0-5 | 0-5 | NA | NA |
| | - Second (PIP2) | 0-5 | 0-5 | 0-4 | 0-4 |
| | - Third (PIP3) | 0-5 | 0-5 | 0-4 | 0-4 |
| | - Fourth (PIP4) | 0-5 | 0-5 | 0-4 | 0-4 |
| | - Fifth (PIP5) | 0-5 | 0-5 | 0-4 | 0-4 |
| | Carpometacarpal (CMC) | | | | |
| | - Third (CMC3) | NA | NA | 0-4 | 0-4 |
| | - Fourth (CMC4) | NA | NA | 0-4 | 0-4 |
| | - Fifth (CMC5) | NA | NA | 0-4 | 0-4 |
| **Wrist** | First metacarpal base (MCB) | 0-5 | 0-5 | NA | NA |
| | Radius bone | 0-5 | 0-5 | NA | NA |
| | Ulna bone | 0-5 | 0-5 | NA | NA |
| | Trapezium/trapezoid (multangular) | 0-5 | 0-5 | NA | NA |
| | Naviculare | 0-5 | 0-5 | NA | NA |
| | Lunate | 0-5 | 0-5 | 0-4 | 0-4 |
| | Multangular navivular-lunate | NA | NA | 0-4 | 0-4 |
| | Radiocarpal | NA | NA | 0-4 | 0-4 |
| **Foot** | Metatarsophalangeal (MTP) | | | | |
| | - First (MTP1) | 0-10 | 0-10 | 0-4 | 0-4 |
| | - Second (MTP2) | 0-10 | 0-10 | 0-4 | 0-4 |
| | - Third (MTP3) | 0-10 | 0-10 | 0-4 | 0-4 |
| | - Fourth (MTP4) | 0-10 | 0-10 | 0-4 | 0-4 |
| | - Fifth (MTP5) | 0-10 | 0-10 | 0-4 | 0-4 |
| | Interphalangeal (IP) | 0-10 | 0-10 | 0-4 | 0-4 |

## 4.10 Ultrasound score

The ultrasound score will be based on ultrasound of 32 joints scored 0-3 for both grey scale (GS) synovitis and power Doppler signal. In addition, two joints and two tendons will be scored but excluded from the ultrasound score. The total GS and PD score will range from 0 to 96, while the total ultrasound score will range from 0 to 192.

**Table 4.5 Overview of the ultrasound scoring**

| Part of the score | Joints | GS left | GS right | PD left | PD right |
|---|---|---|---|---|---|
| **Yes** | Metacarpophalangeal (MCP) | | | | |
| | - First (MCP1) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Second (MCP2) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Third (MCP3) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Fourth (MCP4) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Fifth (MCP5) | 0-3 | 0-3 | 0-3 | 0-3 |
| | Radio-carpal | | | | |
| | Inter-carpal | 0-3 | 0-3 | 0-3 | 0-3 |
| | Radio-ulnar | 0-3 | 0-3 | 0-3 | 0-3 |
| | Elbow | 0-3 | 0-3 | 0-3 | 0-3 |
| | Knee | 0-3 | 0-3 | 0-3 | 0-3 |
| | Talocrural | 0-3 | 0-3 | 0-3 | 0-3 |
| | Metatarsophalangeal (MTP) | | | | |
| | - First (MTP1) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Second (MTP2) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Third (MTP3) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Fourth (MTP4) | 0-3 | 0-3 | 0-3 | 0-3 |
| | - Fifth (MTP5) | 0-3 | 0-3 | 0-3 | 0-3 |
| | Interphalangeal (IP) | 0-3 | 0-3 | 0-3 | 0-3 |

| | | | | | |
|---|---|---|---|---|---|
| ***No*** | *Extensor carpi ulnaris* | *0-3* | *0-3* | *0-3* | *0-3* |
| | *Tibialis posterior tendon* | *0-3* | *0-3* | *0-3* | *0-3* |
| | *PIP2* | *0-3* | *0-3* | *0-3* | *0-3* |
| | *PIP3* | *0-3* | *0-3* | *0-3* | *0-3* |

## 4.11 Magnetic resonance imaging scoring

MRI of the dominant hand and wrist is acquired at baseline, 12, 24 and 36 months. The same side will be assessed at all time-points.

Images will be read according to the OMERACT Rheumatoid Arthritis MRI Scoring (RAMRIS) criteria. The original RAMRIS consists of MRI definitions of important joint pathologies, a core set of MRI sequences and a semi-quantitative scoring system for erosions, joint space narrowing, bone marrow edema, synovitis, and tenosynovitis.

The RAMRIS core set of MRI sequences to assess inflammatory as well as destructive changes in RA joints includes: Imaging in 2 planes, with T1-weighted images before and after intravenous gadolinium-contrast to assess synovitis and erosions; plus a T2-weighted fat saturated sequence or a STIR sequence to assess bone marrow edema.

### 4.11.1 RAMRIS erosion score

Erosion is defined as a sharply marginated bone lesion, with correct juxta-articular localization and typical signal characteristics that is visible in two planes with a cortical break seen in at least one plane. Each bone is scored on a scale of 0-10 based on the proportion of eroded bone compared to the "assessed bone volume" on all available images. For long bones, the "assessed bone volume" is from the articular surface (or its best estimated position, if absent) to a depth of 1 cm, and in carpal bones it is the whole bone.

- No erosion = 0
- 1-10% of bone eroded = 1
- 11-20% of bone eroded = 2
- 21-30% of bone eroded = 3
- 31-40% of bone eroded = 4
- 41-50% of bone eroded = 5
- 51-60% of bone eroded = 6
- 61-70% of bone eroded = 7
- 71-80% of bone eroded = 8
- 81-90% of bone eroded = 9
- 91-100% of bone eroded = 10 (*)
- Joint unable to be scored =U

* When scoring the wrist area, if the bones are fused, score erosions as 10.

The following bones are scored:

Wrist Bones

- Distal radius
- Distal ulna
- Scaphoid
- Lunate
- Triquetrum
- Pisiform
- Trapezium
- Trapezoid
- Capitate
- Hamate
- Proximal metacarpal 1 to 5

MCP Bones

- Proximal proximal phalanx 1 to 5
- Distal metacarpal 1 to 5

Without the first digit (which excludes trapezium and proximal metacarpal 1), the maximum score for the wrist is 130. The maximum score for the MCP bones is 80. With the first digit included (which includes the trapezium and proximal metacarpal 1), the maximum score for the wrist is 150. The maximum score for the MCP bones is 100. The definition of no progression of erosions from baseline to end of study will be an average change in RAMRIS erosion score <1/yearly, with alternative definitions ≤0, <2, <5 units/yearly.

### 4.11.2 Bone Marrow Edema (BME, osteitis)

Bone marrow edema is defined as a lesion, that may occur alone or surrounding an erosion or other bone abnormalities, within the trabecular bone, with ill-defined margins and signal characteristics consistent with increased water contents (high signal intensity on STIR coronal images and low signal intensity on T1 weighted images). Bone marrow edema is assessed in each bone of the wrist and MCP bones. Each bone is scored on a scale of 0-3 based on the proportion of bone with edema.

- No edema = 0
- 1-33% of bone edematous = 1
- 34-66% of bone edematous = 2
- 67-100% of bone edematous = 3
- Joint unable to be scored = U

The following bones are scored:

Wrist

- Distal radius
- Distal ulna
- Scaphoid
- Lunate
- Triquetrum
- Pisiform
- Trapezium
- Trapezoid
- Capitate
- Hamate
- Proximal metacarpal 1 to 5

MCP bones

- Proximal proximal phalanx 1 to 5
- Distal metacarpal 1 to 5

Without the first digit (which excludes trapezium and proximal metacarpal 1), the maximum score for the wrist is 39. The maximum score for the MCP bones is 24. With the first digit is included (which includes trapezium and metacarpal base 1), the maximum score for the wrist is 45. The maximum score for the MCP bones is 30.

### 4.11.3 Synovitis

Synovitis is defined as an area in the synovial compartment that shows above normal post-gadolinium enhancement of a thickness greater than the width of the normal synovium. Enhancement is judged by comparison of T1 weighted images obtained before and after intravenous gadolinium contrast.
The synovitis scale is 0-3. A score of 0 is normal, and 1-3 (mild, moderate, severe) are by thirds of the presumed maximum volume of enhancing tissue in the synovial compartment.

- No synovitis = 0
- Mild synovitis = 1
- Moderate synovitis = 2
- Severe synovitis = 3
- Joint unable to be scored = U

Synovitis is scored 0-3 as described in the EULAR OMERACT RAMRIS reference atlas, with best possible match to the reference images".

The following joints and bones are scored:
Wrist

- Distal Radioulnar
- Radiocarpal
- Intercarpal-Carpometacarpal 2-5
- CMC-1

MCP joints

- MCP-1 to 5

Without the first digit, the maximum score for synovitis for the wrist is 9. The maximum score for the MCP joints is 12. With the first digit included, maximum score for synovitis for the wrist is 12. The maximum score for the MCP joints is 15.

### 4.11.4 Joint Space Narrowing (JSN, Cartilage loss)

JSN is defined as reduced joint space width compared to normal, as assessed in a slice perpendicular to the joint surface. The finding is scored on coronal images at the narrowest point. On coronal T1-weighted images score “white (bone marrow fat) to white”. If surface is eroded, ignore erosions and measure “width if there had been no erosion”. The Cartilage scale is 0-4, with 0 indicating normal to 4 indicating complete ankylosis.

- No Narrowing = 0
- Focal or mild (<33%) narrowing = 1
- Moderate (34% - 66%) narrowing = 2
- Moderate to severe (67% - 99%) narrowing = 3
- Ankylosis = 4
- Joint unable to be scored = U

The following joints are scored:
Wrist

- Radius-scaphoid
- Radius-lunate
- Scaphoid-Lunate
- Lunate-triquetrum
- Scaphoid-trapezium
- Scaphoid-trapezoid
- Capitate-scaphoid
- Capitate-lunate
- Hamate-triquetrum
- Trapezoid-trapezium
- Capitate-trapezoid
- Capitate-hamate
- Carpometacarpal 1 to 5

MCP joints

- Metacarpophalangeal 1 to 5

Without the first digit (excluding trapezium-metacarpal base 1), the maximum score for the wrist is 60. The maximum score for the MCP joints is 16. With the first digit included (including trapezium-metacarpal base 1), maximum score for the wrist is 68. The maximum score for the MCP joints is 20.

### 4.11.5 Tenosynovitis

Tenosynovitis on MRI is defined as tendon sheath fluid, sheath thickening and enhancement after intravenous contrast injection. As small amounts of fluid can be seen in normal tendon sheets, it is essential that the tenosynovitis is visible in at least two consecutive axial slices within the tendon sheet to be scored as abnormal. Tendon sheath abnormalities are graded semi-quantitatively from grade

0 to grade 3, reflecting the maximum width (in mm) of enhancement within each anatomical area as described below:

- Grade 0 (normal): no peritendinous effusion or synovial proliferation with enhancement.
- Grade 1: >0 and <1.5 mm peritendinous effusion and/or synovial proliferation with enhancement.
- Grade 2: ≥ 1.5 mm and < 3 mm peritendinous effusion and/or synovial proliferation with enhancement.
- Grade 3: ≥ 3 mm peritendinous effusion and/or synovial proliferation with enhancement.

The extent of the synovial enhancement is measured at the point of maximal thickness, perpendicular to the tendon surface.

Flexor and extensor tenosynovitis are evaluated semi-quantitatively in 10 different anatomical areas.
Dorsally:

- extensor pollicis brevis, abductor pollicis longus
- extensor carpi radialis brevis, extensor carpi radialis longus
- extensor pollicis longus
- extensor digitorum communis, extensor indicus proprius
- extensor digiti quinti proprius
- extensor carpi ulnaris.

On the volar side:

- the flexor carpi ulnaris tendon (located ulnar to the carpal tunnel)
- the flexor digitorum superficialis and profundus tendons (in the carpal tunnel, enclosed in a common sheath—the ulnar bursa)
- the flexor pollicis longus tendon (located dorsally and radially to the median nerve as it passes through the carpal tunnel, and entering a continuous sheath that becomes the radial bursa)
- the flexor carpi radialis (localised radially to the tendons enclosed in the ulnar bursa)

The maximum tenosynovitis score is 30.

### 4.11.6 Total MRI inflammation score

The total MRI inflammation score will be calculated as the normalized summation of the synovitis score, the osteitis score and the tenosynovitis score, as suggested by the MRI in arthritis OMERACT group (Sundin et al, J Rheum, 2019).

### 4.11.7 Total MRI joint damage score

The total MRI joint damage score will be performed by normalized summation as suggested by the MRI OMERACT group (Sundin et al, J Rheum, 2019).

## 4.12 Remission

Remission status is calculated at each visit. In addition to remission according to cut-offs for the disease activity indices DAS, DAS28, CDAI and SDAI defined previously, the following remission criteria are defined:

### 4.12.1 ACR/EULAR remission

The patient must satisfy all of the following in order to achieve ACR/EULAR remission:

- RAI ≤ 1
- SJC44 ≤ 1
- CRP ≤ 1
- PGA ≤ 1 (on a scale 0-10, in this study ≤ 14 on a scale 0-100)

### 4.12.2 ACR remission

A patient is regarded as in ACR remission if at least 5 of the following criteria are present for at least two consecutive months:

- Morning stiffness ≤ 15 minutes; self-reported
- No fatigue; self-reported fatigue ≤ 14 on VAS (100mm)
- No joint pain; self-reported joint pain ≤ 14 on VAS (100mm)
- No joint tenderness or pain on motion; RAI = 0
- No soft tissue swelling in joints or tendon sheaths; swollen joint count = 0
- ESR ≤ 30 mm/h for female or ≤ 20 mm/h for male

### 4.12.3 FDA remission
Requires achieving ACR clinical remission at the current visit and no radiological progression (change in vdHSS =0) and no use of DMARDs at each visit within the previous 6 months.

### 4.12.4 FDA complete clinical response
Same as FDA remission, but with no requirement on the use of no DMARDs.

### 4.12.5 Complete DAS remission
Based on the FDA's definition of complete clinical response, a complete DAS remission will be defined as: DAS remission (i.e. DAS<1.6) and no swollen joints (i.e. SJC44 = 0) and no radiographic progression (i.e. change in vdHSS =0) for > 6 months.

## 4.13 SF-36
The SF-36 is a multi-purpose, short-form health survey with 36 questions. The SF-36 will be scored according to RAND 36-Item Health Survey 1.0 (http://www.rand.org/health/surveys_tools/mos/mos_core_36item_scoring.html) to form eight measures scores 0-100: physical functioning, bodily pain, role limitations due to physical health problems, role limitations due to personal or emotional problems, emotional well-being, social functioning, energy/fatigue, and general health perceptions. In addition, composite scores for physical and mental health summary measures are calculated according to the New England Medical Centre scoring instructions.(Ware, Kosinski, & Keller, 1994) The composite scores are computed according to the 1998 US general population means and standard deviations.

## 4.14 Physical function
The Stanford Health Assessment Questionnaire was introduced in the 1980s and is now widely used in evaluation of physical function in patients with RA. The disability index of this instrument includes questions concerning the ability of patients to perform 20 activities of daily living, and is most commonly referred to as the HAQ questionnaire, and sometimes as the HAQ disability index (HAQ-DI).

An updated version has been developed, the Patient-Reported Outcomes Measurement Information (**PROMIS**) HAQ, including a 20-item short form used in this study (adult physical function version 1.0 20-item PROMIS short form – 20a). While the original HAQ had 4 response categories, PROMIS includes a fifth response option. Each question has thus five response options, ranging in value from one to five. To find the total raw score, the sum of the values of the response to each question is calculated, giving a range in scores from 20 to 100 if all questions are answered. If at least 50% of the questions are answered, the form can be scored, according to the following formula:

$$\frac{\text{(Raw sum x number of items on the short form)}}{\text{Number of items that were actually answered}}$$

The total raw score should be translated into a T-score for each participant (either by standardized conversion tables or using item-level calibrations), which rescales the raw score into a standardized score with a mean of 50 and a standard deviation (SD) of 10. Therefore, a person with a T-score of 40 is one SD worse than average. The standardized T-score is reported as the final score for each participant.

## 4.15 RAID
The Rheumatoid Arthritis Index of Disease (RAID) is calculated based on seven numerical rating scales (NRS) questions. Each NRS is assessed as a number between 0 and 10. The seven NRS correspond to pain, function, fatigue, sleep, emotional wellbeing, physical wellbeing and coping/self-efficacy.

*Calculation*
RAID final value = (pain NRS value (range 0–10) × 0.21) + (function NRS value (range 0–10) × 0.16) + (fatigue NRS value (range 0–10) × 0.15) + (physical wellbeing NRS value (range 0–10) × 0.12) + (sleep NRS value (range 0–10) × 0.12) + (emotional wellbeing NRS value (range 0–10) × 0.12) + (coping NRS value (range 0–10) × 0.12).
Thus, the range of the final RAID value is 0–10 where higher figures indicate worse status.

*Missing data imputation*
If one of the seven NRS values composing the RAID is missing, the imputation is as follows:
a Calculate the mean value of the six other (non-missing) NRS (range 0–10)
b Impute this value for the missing NRS
c Then, calculate the RAID as explained above.
If two or more of the NRS are missing, the RAID is considered as missing value (no imputation).

### 4.16 EQ-5D
EQ-5D is a standardized generic instrument for use as a measure of health outcome. Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and a single index value for health status. The EQ-5D index values are calculated according to the EQ-5D UK Time Trade-Off (TTO) value set.

### 4.17 WPAI
Worker productivity is generally subdivided into 2 components: absenteeism and presenteeism. The worker productivity in this study is based on the Work Productivity and Activity Impairment Questionnaire: Rheumatoid arthritis V2.0 (WPAI:RA).

The WPAI yields four types of scores:
1. Absenteeism (work time missed)
2. Presenteeism (impairment at work / reduced on-the-job effectiveness)
3. Work productivity loss (overall work impairment / absenteeism plus presenteeism)
4. Activity Impairment

The scores are based on the following questions:

Q1= currently employed
Q2 = hours missed due to specified problem
Q3 = hours missed other reasons
Q4 = hours actually worked
Q5 = degree problem affected productivity while working
Q6 = degree problem affected regular activities

*Scores*
Multiply scores by 100 to express in percentages.
Percent work time missed due to RA (absenteeism):
Percent impairment while working due to RA (presenteeism):

Percent overall work impairment due to RA (work productivity loss):

Percent activity impairment due to problem:

### 4.18 Other calculations
*Age* (years) = [(date of baseline – date of birth)/365.25].

*BMI* = weight in kilograms / (height in metres) x (height in metres)

BMI will be categorized according to the WHO definitions for underweight, normal, overweight and obese.

*Area under the curve* (AUC) will be calculated as the integral under the measure curve using trapezoids

*Time of withdrawal* = date of withdrawal – date of randomization +1

### 4.19 Safety definitions

#### 4.19.1 Treatment emerging adverse events

Treatment emerging adverse events (TEAEs) are defined as AEs with a start date on or after the randomization date.

#### 4.19.2 Past disease and concomitant disease

Past disease/condition

A disease/condition is considered as past disease/condition if it is not ongoing at screening visit.

Concomitant disease

A disease/condition is considered as concomitant disease/condition if it is ongoing at screening visit.

***Previous and Concomitant medications***

- previous medication (start date < date of randomisation);
- concomitant medication (start date ≥ date of randomisation);

In case of missing or incomplete dates/times not directly allowing allocation to any of the two categories of medications, a worst-case allocation was performed according to the available parts of the start and the end dates. The medication was allocated to the first category allowed by the available data, according to the following order:

- concomitant medication
- previous medication

## 5. EFFICACY ENDPOINTS

### 5.1 Primary endpoint

The primary endpoint is the proportion of patients that do not experience a disease flare during the study (as defined for the primary analyses as the first 12 months of follow-up). A disease flare is defined to occur if a patient's ***DAS score increases ≥ 0.6 points to a level in excess of 1.6 (i.e DAS >1.6), and in addition has more than 1 swollen joint on examination of 44 joints.*** If a patient does not fulfill these formal criteria, a disease flare will be recorded if both the patient and investigator agree that a clinically significant flare has occurred.

## 5.2 Secondary endpoints

### 5.2.1 Efficacy endpoints

| Group | Endpoint | Assessment time | Type |
|---|---|---|---|
| **Remission** | DAS remission | All post-baseline visits | Dichotomous, Time to event |
| | DAS28 remission | All post-baseline visits | Dichotomous, Time to event |
| | CDAI remission | All post-baseline visits | Dichotomous, Time to event |
| | SDAI remission | All post-baseline visits | Dichotomous, Time to event |
| | ACR/EULAR remission | All post-baseline visits | Dichotomous, Time to event |
| **Radiology** | ΔvdHSS | Visits 4 | Continuous Δ |
| | ΔErosion score | Visits 4 | Continuous Δ |
| | ΔJSN | Visits 4 | Continuous Δ |
| | Radiographic progression, defined as vdHSS ≥1.0 units/yr | Visits 4 | Dichotomous |
| | Radiographic progression, defined as vdHSS ≥ 0.5 units/yr | Visits 4 | Dichotomous |
| | Radiographic progression defined, as vdHSS ≥ 2.0 units/yr | Visits 4 | Dichotomous |
| | Radiographic progression defined, as vdHSS ≥ 5.0 units/yr | Visits 4 | Dichotomous |
| **Disease Activity** | ΔDAS | All post-baseline visits | Continuous Δ |
| | DAS AUC from baseline | Visits 4 | Continuous AUC |
| | ΔDAS28 | All post-baseline visits | Continuous Δ |
| | DAS28 AUC from baseline | Visits 4 | Continuous AUC |
| | ΔSDAI | All post-baseline visits | Continuous Δ |
| | SDAI AUC from baseline | Visits 4 | Continuous AUC |
| | ΔCDAI | All post-baseline visits | Continuous Δ |
| | CDAI AUC from baseline | Visits 4 | Continuous AUC |
| **ACR core set** | ΔRAI | All post-baseline visits | Continuous Δ |
| | ΔSJC44 | All post-baseline visits | Continuous Δ |
| | ΔPhGA | All post-baseline visits | Continuous Δ |
| | ΔPGA | All post-baseline visits | Continuous Δ |
| | ΔJointPain | All post-baseline visits | Continuous Δ |
| | ΔPROMIS | V4 | Continuous Δ |
| | ΔESR | All post-baseline visits | Continuous Δ |
| | ΔCRP | All post-baseline visits | Continuous Δ |
| **US scores** | ΔGrey scale synovitis | Visit 4 | Continuous Δ |
| | ΔPD synovitis | Visit 4 | Continuous Δ |
| | Joints without PD activity | Visit 4 | Dichotomous |
| **Medication** | Number of patients in DMARD categories | All post-baseline visits | Categorical |

| | | | |
|---|---|---|---|
| | Dose of DMARDs in users | All post-baseline visits | Continuous |
| | Prednisolone usage 0 – 12 months | Visit 4 | Continous |
| | Any intraarticular injections 0 - 12 months | Visit 4 | Dichotomous |
| | Number of intraarticular injections 0-12 months | Visit 4 | Continuous |
| | Total 0-12 month Triamcinolonehexacetonid dose (mg) | Visit 4 | Continuous |
| **Response** | ACR20/50/70/90 | All post-flare visits | Dichotomous, Time to event |
| | FDA major clinical response | All post-flare | Dichotomous |
| | EULAR good response | All post-flare visits | Dichotomous, Time to event |

### 5.2.2 Quality of life endpoints

| Group | Endpoint | Assessment time | Type |
|---|---|---|---|
| **SF-36** | ΔPhysical functioning | V4 | Continuous Δ |
| | ΔBodily pain | V4 | Continuous Δ |
| | ΔRole limitations due to physical health problems | V4 | Continuous Δ |
| | ΔRole limitations due to personal or emotional problems | V4 | Continuous Δ |
| | ΔEmotional well-being | V4 | Continuous Δ |
| | ΔSocial functioning | V4 | Continuous Δ |
| | ΔEnergy/fatigue | V4 | Continuous Δ |
| | ΔGeneral health perception | V4 | Continuous Δ |
| | ΔPhysical health composite score | V4 | Continuous Δ |
| | ΔMental health composite score | V4 | Continuous Δ |
| **RAID** | ΔRAID total score | V4 | Continuous Δ |
| **EQ5D** | ΔEQ5D index value | V4 | Continuous Δ |
| **WPAI** | ΔAbsenteeism | V4 | Continuous Δ |
| | ΔPresenteeism | V4 | Continuous Δ |
| | ΔWork productivity loss | V4 | Continuous Δ |
| | ΔActivity impairment | V4 | Continuous Δ |

# 6. SAFETY PARAMETERS

Measures of safety will include the following:

- Clinical and laboratory adverse events (AEs) and coding of AEs performed (using the [Medical Dictionary for Regulatory Activities] MedDRA, v.21.1E).
- Clinical laboratory data
- Vital signs

# 7. STATISTICAL METHODOLOGY

## 7.1 Statistical and Analytical Issues

### 7.1.1 Statistical Methods

The primary efficacy analyses will be based on the PPS. Secondary efficacy analyses will be carried out in the PPS populations, with robustness analyses of primary and secondary efficacy analyses in the FAS population. Baseline characteristics will be primarily be performed in the FAS population, with supplementary description of the PPS population. There is only one primary endpoint and thus only one primary analysis, and multiple testing adjustments will not be made in the secondary analyses.

All categorical (including binary and ordinal) data will be summarized using frequency counts and percentages of patient incidence. Percentages will be calculated using the appropriate study population (FAS or PPS); any exceptions to this will be highlighted in the table footnote. The continuous variables will be summarized using number of patients (N), mean, standard deviation (SD), median, 25/75 percentile and range (minimum/maximum).

All efficacy analyses will be presented by the size (point estimate) of the difference between the treatments and the associated 95% confidence interval.

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

#### *7.1.1.1 Analyses of dichotomous endpoints*

Dichotomous efficacy endpoints (including the primary and secondary endpoints of section 5.2.1) at 12 months (visit 4) will be analyzed using mixed effect logistic regression with treatment group as the only fixed effect and center as a random effect.

Specifically, let $y_{i,c}$ (=1 if non-failure, 0 otherwise) denote the response for the $i$-th subject ($i = 1,..,n_c$) from the $c$-th center ($c = 1,..,C$), with treatment group $x_i$ (=1 if in stable sDMARD therapy, 0 otherwise). The analysis model is then

$$y_{i,c} \sim Bernoulli(p(x_i, c))$$

where $p(x,c) = (1 + e^{-(\beta_0+\beta_1 x+\alpha_c)})^{-1}$. The center effect is assumed to be normally distributed, i.e. $\alpha_c \sim N(0, \sigma^2)$, with independence between centers.

The model parameters will be estimated using maximum likelihood estimation. The treatment difference will be estimated as the marginal mean for stable sDMARD treatment minus the marginal mean for tapered sDMARD treatment. The confidence interval for the treatment difference will be based on the delta method.

Efficacy assessments of dichotomous end-points through time will also be analyzed with a mixed effect logistic regression model. Here, in addition to the center effect, subject specific random intercepts will be included. The fixed effects will be the factors: treatment and visit, including a treatment-by-visit interaction. Treatment differences at different time points will be assessed using marginal means as above.

#### *7.1.1.2 Analyses of continuous endpoints*

Continuous endpoints will be analyzed using the linear mixed effect model. Each analysis will include an adjustment for the baseline value of the endpoint, and further include treatment, visit, and their interaction as fixed factors. Subject and center specific intercepts will be treated as random effects, and differences between treatment groups will be estimated using marginal means.

#### *7.1.1.3 Analyses of time to event endpoints*

Time to event endpoints (see section 5.2.1) will be analyzed using a Weibull regression model with center as a random effect. As above, the random effects will be assumed to be normally distributed. Sensitivity analyses will be performed using Cox regression and Kaplan-Meier product-limit analysis. Estimates of the hazard ratio will be presented in addition to Kaplan-Meier plots.

#### *7.1.1.4 Software implementation*

All analysis will be done using Stata v14. Estimation of the logistic mixed effect models will be done using the *melogit* function, the linear mixed effect models via the *mixed* function, and the time-to-event endpoints analyzed via the *mestreg* for function for multi-level survival analysis.

The primary end-point treatment difference will be assessed using that average marginal effect, estimated by the *margins* function. Combined with *melogit*, this estimates the marginal effect, integrating over the unconditional distribution of random effects.

The stata margins command fails if a variance component is estimated too close to zero. For this reason, if this occurs, the model is re-estimated with the corresponding random effect removed, and the margins command run on the reduced model.

## 7.2 Determination of sample size

To estimate the expected percentage of patients who will be non-failures after one year of follow-up in the ARCTIC REWIND studies, we used data from the Norwegian NOR-DMARD register, with TNFi treatment as the example. Patients who had one year of sustained DAS28 remission and were receiving TNF inhibitors in combination with MTX were selected. Of these, 80% were still in remission one year later. The corresponding number for patients receiving MTX monotherapy was 74%. Based on this, we hypothesize that also 80% of the patients included in ARCTIC REWIND would be in remission (an approximation of non-failure) after one year. An inferiority margin of 20% was chosen, as it was thought that clinicians would accept an 20% increased risk of flare given the potential benefits of decreasing DMARD therapy.

If there is truly no difference between half dose and continued full dose DMARD treatment on the proportion of patients in remission after 12 months, 126 patients (63 in each arm) are required in to be 80% confident that the upper limit of a one-sided 97.5% confidence interval (equivalent to a 95% two-sided confidence interval) will exclude a more than 20% difference in favor of the full-dose treatment. Different combinations of remission rates and non-inferiority margins are summarized in the Table 4.

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 160 patients will be randomized for the sDMARD study.

**Table 4:** The numbers in the cells represent the total number of patients needed under different scenarios. All calculations are based on a power of 80% and alpha 2.5%.

| Non-inferiority Margin | 90% remission at 12 m | 80% remission at 12 m | 70% remission at 12m |
|---|---|---|---|
| 15 % | 126 | 224 | 294 |
| 20 % | 72 | 126 | 166 |
| 25 % | 46 | 82 | 106 |

## 7.3 Handling of Dropouts and Missing Data

In general, missing values will not be imputed for descriptive statistics.

### 7.3.1 Primary endpoint

For the analysis of the primary endpoint, consisting of flare assessments, the patients in the PPS will not have missing values for the primary endpoint, as the primary variable has been fully monitored during the study and no values are missing. For analyses in the FAS population, patients with more than one missing visit will be set to worst outcome (flare).

### 7.3.2 Other dichotomous endpoints

All other dichotomous endpoints analyzed via logistic mixed models using all assessments. Missing values will be assumed to be missing at random and will not be imputed.

### 7.3.3 Other continuous endpoints
All continuous endpoints will be analyzed using the linear mixed model and no imputation will be performed. Missing values will be assumed missing at random.

### 7.3.4 Time to event endpoints
For the time to event analysis, all patients who withdraw from follow-up will be censored on the withdrawal date.

### 7.3.5 ACR20, ACR50 and ACR70
Because the ACR20, ACR50, ACR70 variables are based on several component variables, it is possible that the values may still be calculated even if the component variables have some missing values. In this case, no imputation method is needed. If the ACR value is still missing, an imputation method will be applied. If the ACR value is missing due to missing values in any of the components, while the patient is still enrolled, the method of last observation carried forward (LOCF) will be used to carry forward any of the missing components, and from that mix of actual and carried-forward values, the values of ACR20, ACR50 and ACR70 will be determined.

After the LOCF imputation has been applied, missing values due to a patient dropping from the study for any reason (e.g. lack of efficacy or adverse event) will be handled by setting the ACR value (ACR20, ACR50 and ACR70) to nonresponsive (that is, observation at time of flare carried forward) from that visit onward.

### 7.3.6 Pooling of Investigator Sites
There will be no pooling of investigator sites. Handling of centre-effects is described previously.

### 7.3.7 Timing of Main Analysis
The analysis of primary endpoint and main secondary endpoints during the first 12 months of the study is planned to occur when all eligible patients have completed 12 months of treatment, all data up to 12 months have been entered, verified and validated and the primary database has been locked.

Analysis on MRI data (RAMRIS-scores) is planned when all images are received, scored and entered into the MRI part of the database. This will be done after the main analysis.

Analysis of the last study periods (study period 2 and 3, i.e. 1-2 years and 2-3 years) will be formed after all eligble patients have completed the study periods, all data for the study periods have been entered, verified and validated and the database has been locked.

## 7.4 Patient Characteristics

### 7.4.1 Patient Disposition
The disposition of all patients will be listed and summarised by treatment arm. The number and percentage of patients who are randomised, received any study treatment, prematurely discontinued from treatment and lost to follow-up will be summarized (supplementary figure 1).

The number and percentage of patients will be categorized by the reason(s) for

1) End of study treatment/withdrawal from study treatment: This is when it is decided that a patient will not receive further treatment according to study protocol and enters the survival follow-up phase. Reasons can be: adverse event, patient withdrawal of consent, investigator decision, death, lost to follow-up, wrong diagnosis, major protocol deviation, unknown, other.

### 7.4.2 Protocol Deviations
Major protocol deviators will be determined and summarised by type (Major/Minor) and treatment group.

Major protocol deviations are:
- Informed consent not dated and signed
- Eligibility, according to Inclusion / Exclusion criteria, not met
- Serious adverse event (SAE) reporting requirements not met; seriousness criteria misinterpreted, timelines not respected
- Treatment regimen deviation: if the patient for some reason does not comply to the treatment strategy defined in the protocol
- Randomisation non-compliance; patient starts other treatment strategy than allocated

### 7.4.3 Background and Demographic Characteristics
Patient demographics and baseline characteristics will be summarized for the FAS population.

Patient demographics and baseline characteristics will be summarized by randomized 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. The patient demographics and baseline characteristics to be summarized include age in years, sex, symptom duration, anti-CCP status, rheumatoid factor status, BMI, work status, smoking status, CRP, TJC, SJC, RAI, physician global VAS, patient global VAS, patient VAS pain, fatigue VAS, DAS, DAS28, CDAI, SDAI, ACR/EULAR Boolean remission, EQ-5D, SF-36 mental and physical component summary scores, van der Heijde modified Sharp scores, HAQ, ultrasound scores and medication.

Medical history will be coded using the MedDRA dictionary (v21.1E) and will be summarised. Concomitant medication will be coded using the ATC coding system and summarised.

### 7.4.4 Baseline DMARD treatment
Baseline DMARD treatment will be summarised by categories outlined in table 7.2.

**Table 7.1 Treatment regimen**

| Description |
|---|
| Methotrexate monotherapy |
| Sulfasalazine monotherapy |
| Leflunomide monotherapy |
| Methotrexate + sulfasalazine |
| Methotrexate + hydroxychloroquine |
| Methotrexate + hydroxychloroquine + sulfasalazine |

### 7.4.5 Treatment Compliance
Data summarizing the proportions of patients complying with the treatment regimen according to protocol will be analyzed by treatment arm. Treatment arm difference will be tested using the chi-square test.

### 7.4.6 Concomitant Medications and Other Therapies
Concomitant medication information collected will be coded by the Anatomical Therapeutic Chemical (ATC) classification system. Concomitant medications taken during the study (including any prior medication that is continuing at the time of inclusion) will be classified by generic name. The number and percentage of patients who took at least one drug within each specific preferred term will be assessed. Patients will only be counted once if they are taking more than one medication (within the same code) or take the same generic medication more than once. If it cannot be determined whether a medication is concomitant (based on stop date or, if the stop date is missing, start date), then the medication will be considered to be concomitant.

### 7.4.7 Quality of Life data

Analyses of Quality of Life (QoL) data will be done using the procedures described for change from baseline continuous endpoints (section 7.1.1.2).

## 7.5 Disease activity before, under and after flares

For patients who experience flares (Flare Analysis Set), we will assess the disease activity at the visit prior to the flare, at the flare visit, and at subsequent visits in the two arms.

## 7.6 Identification of predictors for flare

Two approaches to identify predictors for flares will be explored, based on two populations:
TAS population: assessment of the risk of disease flare under tapered therapy

PPS population: assessment of the risk of disease flare under stable and tapered therapy regimens

In the TAS population, a model for predicting disease flare will be based on baseline characteristics (e.g age, sex, smoking status, BMI, symptom duration, disease activity (index based, fulfillment of different remission criteria, individual variables from the composite scores), serological status (positivity, levels), ultrasound findings, radiographic scores, DMARD regimen (including MTX monotherapy p.o. vs MTX monotherapy s.c.)). First, the univariate association between each of the baseline characteristics and any disease flaring during follow-up will be assessed using separate logistic regression models. Second, a multivariable predictive model will be formed using step-wise selection. For an unbiased assessment of the predictive accuracy of the model the Area Under the Curve of the model (and its 95% CI) will be assessed by 10-fold cross-validation, at each fold repeating the model selection.

In the PPS population, a similar model will be developed using data from both treatment groups. Here the step-wise model selection will be applied to the same set of baseline disease characteristics, but in addition including their interaction with treatment group. From the final model the risk difference of disease flaring under tapered and stable therapy will be formed for each patient along with its 95% confidence interval. The confidence interval for the patient level risk difference will be formed using bootstrapping applied to the model selection and estimation process.

## 7.7 Exploratory Analysis

Samples (including serum, plasma, full blood and urine) for biomarker or DNA/RNA discovery and validation have been collected and stored in a freezer at -70 C. These samples will be used for exploratory analyses, and may include measurement of cytokines and other known or potential new markers of inflammation or damage, such as interleukins, interferons, metalloproteases, transforming growth factor, TNFs, adhesion molecules etc. as well as DNA/RNA analyses (genomics and proteomics).

Exploratory endpoints will not be limited to those mentioned above, and will include variables/endpoints and statistical methods/modelling as necessary to explore the secondary objectives of the study as described in section 2.2 of the protocol.

# 8. SAFETY ANALYSIS

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 and clinical laboratory results. Safety variables will be tabulated and presented for all patients in the safety set.

## 8.1 Adverse Events

Adverse events will be coded using MedDRA, version 21.1E. 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.

The number (%) of subjects with any adverse event, with 1, 2 or $\geq$ 3 adverse events, with treatment related adverse events, and with SAE will be summarized 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 summarized by treatment group, overall, for severe AEs and for AEs leading to study discontinuation. In addition, a summary table of AEs reported by >= 5% of all patients will be presented by SOC and PT. A detailed patient narrative will be given for any death or cancer, in addition to other relevant serious adverse events, in the clinical study report in addition to listing.

### 8.1.1 Clinical Laboratory Parameters

Safety clinical laboratory parameters were collected and assessed, but only used to identify adverse events. No analyses of clinical laboratory parameters will be performed.

### 8.1.2 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 assessment time and treatment arm.

## 8.2 Interim Analyses

There are no planned interim analyses for efficacy.

### 8.2.1 Independent Data Monitoring Committee (IDMC)

No analyses have been made for the IDMC, and will not be reported in the clinical study report.

## 9. DATA ANALYSES FOLDER PLAN (RD)

All programs and datasets will be organized according to the following plan.

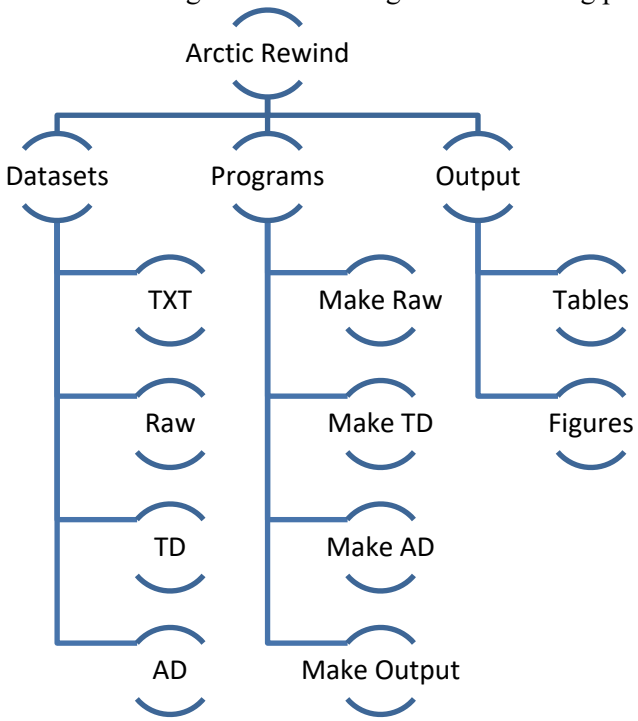


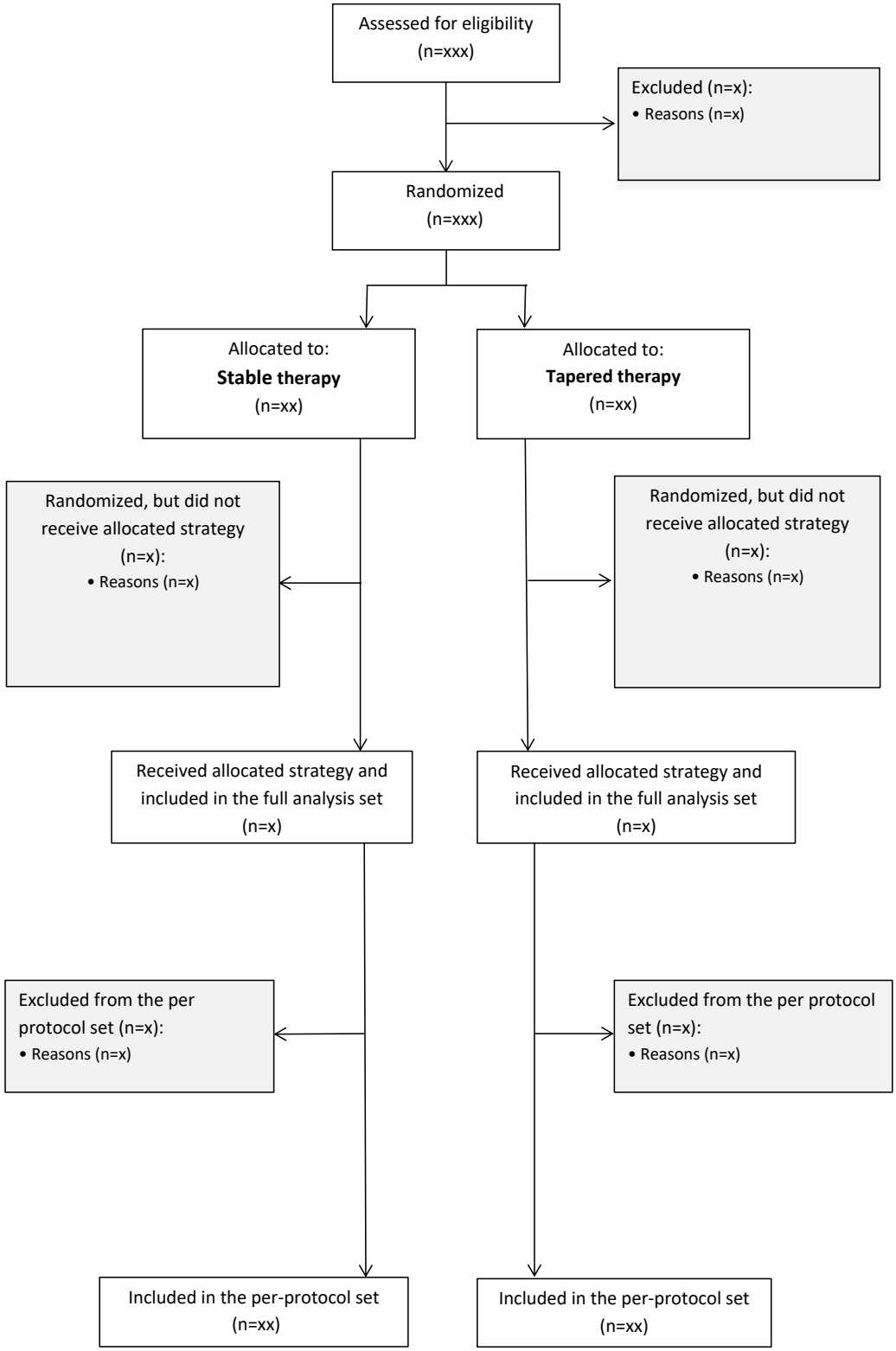

**Figure 9.1 Data analysis folder plan**

- Datasets
  - TXT; all exported data from the eCRF (Viedoc)
  - Raw; all exported files converted to Stata formal
  - TD; Tabulated datasets formed by 'Make TD' scripts. All TD datasets are created from raw datasets
  - AD; Analysis datasets formed from TD data using scripts in 'Make AD'
- Programs
  - Make Raw: Programs for importing TXT-data into Stata format, including variable naming an labeling
  - Make TD: Programs for combining and restructuring Raw-data files.
  - Make AD: Programs combining TD-data into datasets ready for analysis .
  - Make Output: Programs for analyzing data, creating tables and figures.
- Output
  - Tables: Results of analysis in Table form.
  - Figures: Results of analysis in Figure form.

## 10. PLANNED TABLES AND FIGURES

Tentative figures and tables are outlined on the following pages, and will be updated as necessary to comply with journal requirements, including addition of extra data or reduction in size as necessary. Supplemental figures and tables will be included in accordance with journal guidelines.

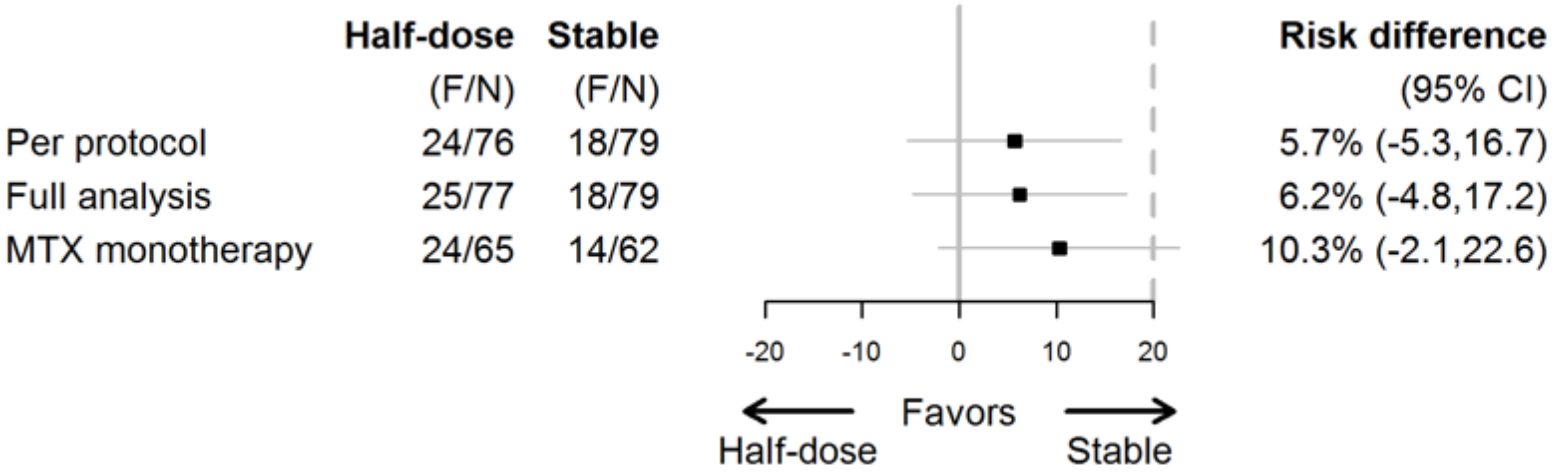
**Figure 1: Patient disposition**



| **Table 1. Characteristics of patients at baseline.*** | | |
|---|---|---|
| **Characteristic** | **Stable arm (N=)** | **Tapered arm (N=)** |
| Age – yr | | |
| Female sex – no. (%) | | |
| Body-mass index (kg/m$^2$) | | |
| Current smoker – no. (%) | | |
| Time since first swollen joint – years | | |
| Positive for anti-citrullinated peptide antibodies – no. (%) | | |
| Positive for rheumatoid factor – no. (%) | | |

* Plus – minus values are means ± SD. Median values are given with interquartile range (IQR). There were no significant differences between the groups except for

**Figure 2: Non-inferiority plot of stable vs tapered sDMARD treatment in per protocol set, full analysis set and in patients treated by methotrexate monotherapy**.
*(Illustration based on dummy variables)*

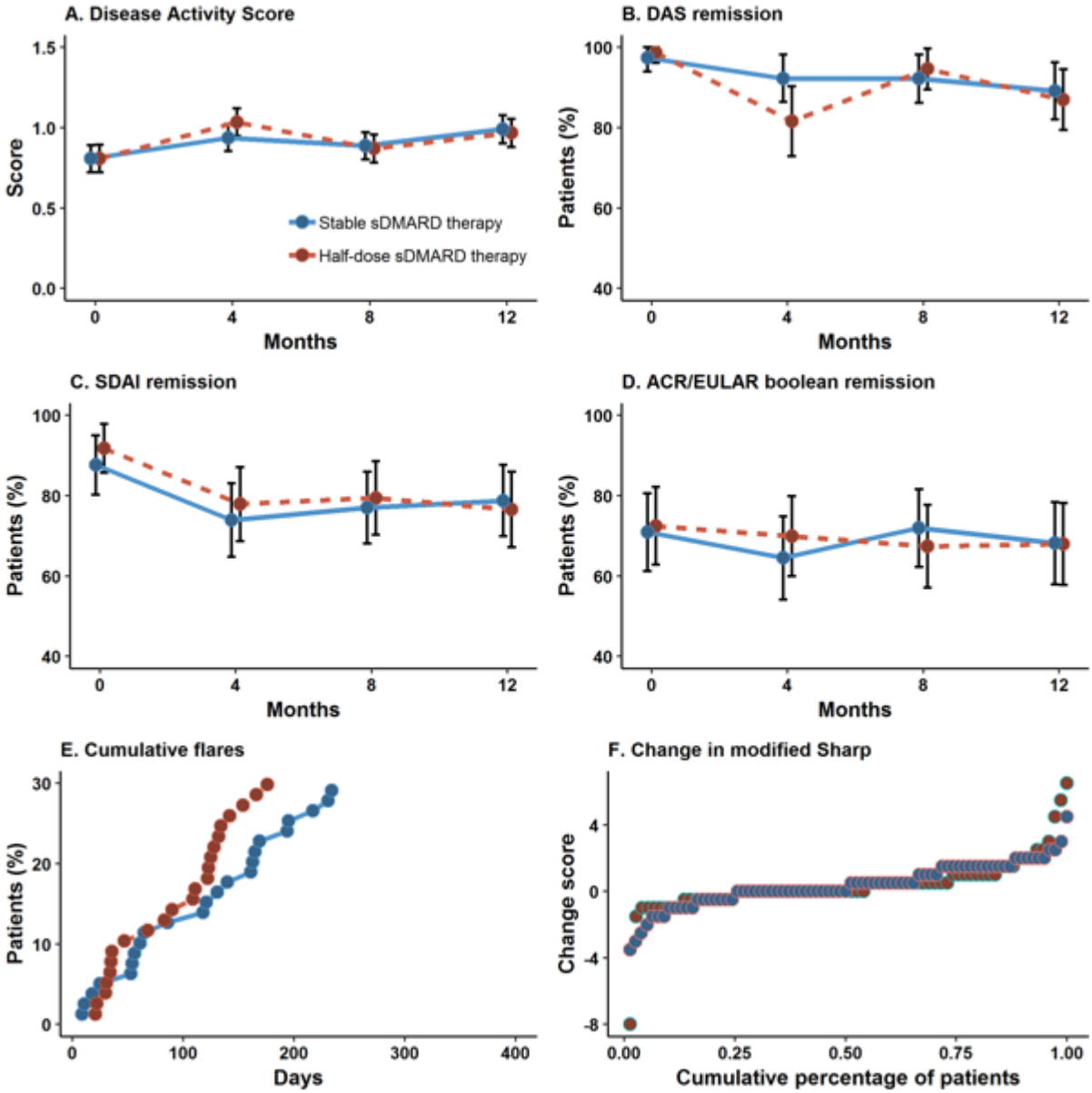


| Table 2. Secondary outcomes.* | | | | | |
|---|---|---|---|---|---|
| | Baseline | | 12 months | | Difference at 12 months (95% CI) |
| | Stable arm | Tapered arm | Stable arm | Tapered arm | |
| **Continous variables** | **Observed values** | | **Change from baseline** | | |
| **Measures of disease activity** | | | | | |
| Disease Activity Score | | | | | |
| Disease Activity Score in 28 joints | | | | | |
| Simplified Disease Activity Index | | | | | |
| Clinical Disease Activity Index | | | | | |
| Swollen-joint count | | | | | |
| Tender-joint count (Ritchie Articular Index) | | | | | |
| Erythrocyte sedimentation rate, mm/hr | | | | | |
| C-reactive protein, mg/liter | | | | | |
| Patient's global assessment | | | | | |
| Investigator's global assessment | | | | | |
| **Functional outcomes** | | | | | |
| PROMIS Physical Function | | | | | |
| EuroQol-5 Dimensions | | | | | |
| Fatigue visual-analogue scale | | | | | |
| SF-36 Physical Component Summary Score | | | | | |
| SF-36 Mental Component Summary Score | | | | | |
| WPAI % work missed due to specified problem (absenteeism) | | | | | |
| WPAI % impairment while working due to specified problem (presenteeism) | | | | | |
| WPAI % overall work impairment due to specified problem | | | | | |
| WPAI % activity impairment due to specified problem | | | | | |
| Rheumatoid Arthritis Impact of Disease total score | | | | | |
| **Radiographic joint damage** | | | | | |
| total van der Heijde modified Sharp score | | | | | |
| *van der Heijde Sharp Erosion* | | | | | |
| *van der Heijde Sharp Joint Space Narrowing* | | | | | |
| **Ultrasound outcomes** | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Total power Doppler signal score | | | | | |
| Total grey scale score | | | | | |
| **Medication** | | | | | |
| Dose methotrexate in users | | | | | |
| Dose sulfasalazine in users | | | | | |
| Dose hydroxychlorochine in users | | | | | |
| Dose leflunomide in users | | | | | |
| Total Triamcinolonehexacetonid dose (mg) | **NA** | **NA** | | | |
| Number of intraarticular injections | **NA** | **NA** | | | |
| **Cathegorized variables** | **Observed values** | | **Observed values** | | |
| **Measures of diseaes activity** | | | | | |
| Disease Activity Score remission – no. (%) | | | | | |
| Disease Activity Score in 28 joints remission – no.(%) | | | | | |
| Clinical Disease Activity Index remission – no. (%) | | | | | |
| ACR/EULAR remission – no. (%) | | | | | |
| No swollen joints – no. (%) | | | | | |
| **Imaging outcomes** | | | | | |
| No radiographic progression – no. (%) | | | | | |
| No power Doppler signal in any joint – no. (%) | | | | | |
| **Medication** | | | | | |
| Methotrexate monotherapy – no. (%) | | | | | |
| Methotrexate monotherapy p.o. – no. (%) | | | | | |
| Methotrexate monotherapy s.c. – no. (%) | | | | | |
| Sulfasalazine monotherapy – no. (%) | | | | | |
| Leflunomide monotherapy – no. (%) | | | | | |
| Methotrexate + sulfasalazine – no. (%) | | | | | |
| Methotrexate + hydroxychloroquine – no. (%) | | | | | |
| Methotrexate/sulfasalazine/hydroxychloroquine – no. (%) | | | | | |
| Biologic treatment – no. (%) | NA | NA | | | |
| Any intraarticular injections 0-12 months– no. (%) | NA | NA | | | |
| Any prednisolone use over 12 months – no. (%) | NA | NA | | | |

* Plus – minus values are means ± SD. Median values are given with interquartile range (IQR). There were no significant differences between the groups except for

**Figure 3 – Secondary endpoints.** Panel A shows the Disease Activity Score. Panel B shows the percentage of patients in Disease Activity Score remission. Panel C shows the percentage of patients in Simplified Disease Activity Index remission. Panel D shows the percentage of patients in ACR/EULAR Boolean remission. Panel E shows the cumulative percentage of patients who have experienced flares. Panel F shows the cumulative probability plot for radiographic joint damage, scored by van der Heijde Sharp score. *(illustration based on dummy variables)*



| **Table 3:** Safety findings from month 0 to 12. | | |
|---|---|---|
| **Event** | **Stable arm (N=xx)** | **Tapered arm (N=xx)** |
| Any adverse event – no. (%) | | |
| Patients with adverse events | | |
| 1 adverse event – no. (%) | | |
| 2 adverse events – no. (%) | | |
| ≥3 adverse events – no. (%) | | |
| Serious adverse events – no. (%) | | |
| Adverse events of special interest | | |
| Infection – no. (%) | | |
| Serious infection – no. (%) | | |
| Cancer – no. (%) | | |
| Death – no. (%) | | |
| Adverse event leading to study discontinuation – no. (%) | | |

| **Suppl table: Safety data from 0 to 12 month. Values are number (percentages)** | | |
|---|---|---|
| **MedDRA System Organ Class** | **Stable arm** | **Tapered arm** |
| **All-Cause Mortality***– no. (%) | | |
| | | |
| | | |
| **All Serious Adverse Events***– no. (%) | | |
| | | |
| **Other Adverse Events#** | | |
| Gastrointestinal disorders – no. (%) | | |
| Infections and infestations – no. (%) | | |
| Injury, poisoning and procedural complications – no. (%) | | |
| Musculoskeletal and connective tissue disorder – no. (%) | | |
| Renal and urinary disorders – no. (%) | | |
| Respiratory, thoracic and mediastinal disorder – no. (%) | | |
| Surgical and medical procedures – no. (%) | | |
| Neoplasms beningn, malignant and unspecified (incl cysts and polyps) – no. (%) | | |
| Nervous system disorder – no. (%) | | |
| Cardiac disorder – no. (%) | | |
| Skin and subcutaneous tissue disorders – no. (%) | | |
| General disorders and administration site conditions – no. (%) | | |
| Eye disorder – no. (%) | | |
| Hepatobiliary disorders – no. (%) | | |
| Immune system disorders – no. (%) | | |
| Investigations – no. (%) | | |
| Blood and lymphatic system disorders – no. (%) | | |
| Vascular disorders – no. (%) | | |
| Ear and labyrinth disorder – no. (%) | | |
| Psychiatric disorders – no. (%) | | |
| Reproductive system and breast disorders – no. (%) | | |

## 11. REFERENCES

Hammer, H. B., Bolton-King, P., Bakkeheim, V., Berg, T. H., Sundt, E., Kongtorp, A. K., & Haavardsholm, E. A. (2011). Examination of intra and interrater reliability with a new ultrasonographic reference atlas for scoring of synovitis in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, *70*(11), 1995–1998. http://doi.org/10.1136/ard.2011.152926

Reeve, R., Pang, L., Ferguson, B., O'Kelly, M., Berry, S., & Xiao, W. (2013). Rheumatoid Arthritis Disease Progression Modeling. *Therapeutic Innovation & Regulatory Science*, *47*(6), 641–650. http://doi.org/10.1177/2168479013499571

Ware, J. E., Kosinski, M., & Keller, S. D. (1994). SF-36 physical and mental health summary scales. Boston, MA.