

Disease management in early rheumatoid arthritis

Development and implementation of novel treatment strategies

**Thesis by
Anna-Birgitte Aga**

2016



**Diakonhjemmet Hospital
Department of Rheumatology
Oslo, Norway**



**University of Oslo
Faculty of Medicine
Oslo, Norway**

© Anna-Birgitte Aga, 2016

*Series of dissertations submitted to the
Faculty of Medicine, University of Oslo*

ISBN 978-82-8333-285-8

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.
Print production: Repräsentralen, University of Oslo.

Contents

Acknowledgements	4
Funding	6
Abbreviations	7
List of papers	9
1. Introduction and background	11
1.1 Rheumatoid arthritis	11
1.1.1 Classification	11
1.1.2 Epidemiology	11
1.1.3 Pathogenesis	13
1.1.4 Treatment	13
1.1.5 Treatment strategies	15
1.2 Assessment of disease activity and physical function	16
1.2.1 Single outcome measures	16
1.2.2 Composite outcome measures	17
1.2.3 Remission criteria	18
1.2.4 Biochemical and immunological markers	19
1.2.5 Physical function and health-related quality of life	19
1.3 Imaging in rheumatoid arthritis	20
1.3.1 Conventional radiography	21
1.3.2 Ultrasound	22
1.3.3 Magnetic resonance imaging	24
1.3.4 Imaging remission in rheumatoid arthritis	24
2. General aim and research questions	27
2.1 General aim	27
2.2 Research questions	27
3. Material and methods	28
3.1 Study design	28
3.2 Study populations and organization of data collection	28
3.2.1 The NOR-DMARD study	28
3.2.2 The ULRABIT study	29

3.2.3 The ARCTIC trial.....	30
3.3 Data collection.....	36
3.3.1 Demographic and other background variables.....	36
3.3.2 Clinical, serological and physical outcome measures.....	36
3.3.3 Ultrasound.....	38
3.3.4 Conventional radiographs.....	41
3.3.5 Adverse events and discontinuations.....	41
3.4 Statistics.....	41
3.4.1 Descriptive statistics.....	41
3.4.2 Group comparison.....	42
3.4.3 Multivariate analyses.....	42
3.4.4 Survival analysis.....	42
3.4.5 Factor analysis.....	42
3.4.6 Responsiveness.....	43
3.4.7 Cumulative probability plot.....	43
3.4.8 Sample size estimation in the ARCTIC trial.....	43
3.4.9 Hypotheses and decision rules in the ARCTIC trial.....	44
3.4.10 Analyses sets in the ARCTIC trial.....	44
3.4.11 Missing data.....	45
3.5 Legal and ethical aspects.....	46
3.6 Good clinical practice and monitoring.....	46
4. Summary of results.....	47
4.1 Paper I.....	47
4.2 Paper II.....	48
4.3 Paper III.....	49
4.4 Paper IV.....	50
5. Discussion.....	51
5.1 Discussion of methodological aspects.....	51
5.1.1 Study design.....	51
5.1.2 Representativeness of study populations.....	52
5.1.4 Statistical considerations.....	54
5.2 Discussion of main results.....	55
5.2.1 Baseline disease activity level and disease duration 2000–2010.....	55
5.2.2 Response and remission rates 2000–2010.....	56

5.2.3 Response and remission rates 2010–2015	58
5.2.4 Feasibility of ultrasound remission as treatment target in early rheumatoid arthritis	59
5.2.5 Effect of ultrasound guided treatment of early rheumatoid arthritis	60
5.2.6 Information retained in a reduced ultrasound inflammation score.....	61
5.2.7 Responsiveness of the ultrasound score in paper IV and comparison to former scores....	62
6. Conclusions	64
6.1 Answer to research questions	64
6.2 Clinical implications	65
7. References	67
8. Papers I-IV	87

Acknowledgements

This thesis could not have been possible without the contribution of a great number of people to whom I am incredibly grateful.

First of all, I would like to thank my main supervisor Professor Espen A. Haavardsholm for introducing me to the world of research, for sharing his wide knowledge, for his patience and support, for believing in me, and giving me responsibility. His ability to plan and carry out large randomised trials is impressive, and I am thankful for being included in the ARCTIC-trial from the beginning. Most of all I want to thank him for being a good friend.

I also want to thank Professor Tore K. Kvien. Firstly, for the work he has put into the department over years, creating the stimulating work environment that it is today. I would like to thank him for employing me as a young fellow in the clinical department in 2003, and further inviting me into the academic world of rheumatology after several years of clinical practice.

I would like to thank my other co-supervisors. Elisabeth Lie – for introducing me to the NOR-DMARD study and statistics, for discussing research questions and results with me for hours, and introducing me to scientific writing in an excellent way. Professor Till Uhlig – for always being supportive and available for discussions when needed.

I am also grateful to Hilde B. Hammer for sharing her knowledge and enthusiasm for ultrasound in rheumatology. I owe my gratitude to Professor Désirée van der Heijde, who has contributed with her impressive amount of knowledge and methodologically exceptional mind. I am thankful to Siri Lillegraven for discussions, great support and encouragement day by day, and her qualified advice regarding methods and writing. Thanks to Inge Christoffer Olsen for statistical lectures, discussions on statistical methods and valuable advice.

There would have been no research without the participating patients. Thank you all for spending time on examinations and questionnaires.

I would like to express my thanks to the local investigators in the ARCTIC trial and the NOR-DMARD study, who are my co-authors; Hallvard Fremstad,

Tor Magne Madland, Åse Stavland Lexberg, Hilde Haukeland, Erik Rødevand, Christian Høili, Hilde Stray, Anne Lindtner Noraas , Inger Johanne Widding Hansen, Gunnstein Bakland, Ada Wierød, Knut Mikkelsen, Synøve Kalstad and Dag Magnar Soldal, and research nurses and clinicians in their respective departments in Ålesund, Bergen, Drammen, Martina Hansen Hospital, Trondheim, Moss, Haugesund, Kristiansand, Tromsø and Lillehammer.

I am very grateful for being surrounded by my research colleagues, both current PhD candidates and Postdocs in the “White House”. I am thankful for lunches and coffee breaks with scientific chats or very non-scientific discussions. Special thanks to Inger Jorid, Karen and Eirik, whom I have shared office with.

I would also like to offer thanks to my colleagues in the clinical work, doctors, nurses and secretaries. They are reminding me of having the patients and day-to-day clinical decision-making in mind when doing research. Special thanks to Ellen and Anne Katrine for collaboration in the out-patient clinic.

I am thankful to Diakonhjemmet Hospital for the excellent support they provide their researchers with, and the Norwegian Women’s Public Health Association for founding.

Last, but not least, I would like to thank my family and friends. Special thanks to my father Olav, my family-in-law and my sisters Maud-Kristine and Elisabeth for endless care and support. The final thanks goes to my husband, Stian, for his love, patience and support, and our three wonderful sons Osmund, Oscar and Odd-Olav who are reminding me of what the most important things in life are.

Funding

The PhD-fellowship was granted by the Norwegian Women's Public Health Association and institutional support provided by the administration of Diakonhjemmet Hospital.

The ARCTIC trial was supported by the Norwegian Research Council, the South-Eastern Norway Regional Health Authority, the Norwegian Rheumatism Association, and investigator initiated research grants from AbbVie, UCB, Pfizer, MSD and Roche. Siemens Healthcare and GE Healthcare provided technical support regarding standardization of ultrasound equipment.

The NOR-DMARD study has received financial support from the Norwegian Directorate of Health and Social Affairs and the following pharmaceutical companies: Abbott/AbbVie, Amgen, Aventis, Bristol-Myer Squibb, Roche, Schering-Plough/Centocor/MSD, UCB and Wyeth/Pfizer.

The ULRABIT study has received research grants from AbbVie, Roche, Pfizer, Bristol-Myer Squibb and UCB.

The funders of the studies have not been involved in study design, data collection, data analyses, data interpretation, or writing of the manuscripts.

Abbreviations

ACPA	Anti-Citrullinated Peptide Antibodies
ACR	American College of Rheumatology
AE	Adverse Event
Anti-CCP	Antibodies to Cyclic Citrullinated Peptide
ARCTIC	Aiming for Remission in rheumatoid arthritis: a randomised trial examining the benefit of ultrasonography in a Clinical Tight Control regimen
CDAI	Clinical Disease Activity Index
CI	Confidence Interval
CR	Conventional Radiography
CT	Computed Tomography
CRP	C-Reactive Protein
DAS	Disease Activity Score
DIP	Distal InterPhalangeal
DMARD	Disease Modifying AntiRheumatic Drug
eCRF	Electronic Case Report Form
ECU	Extensor Carpi Ulnaris
ESR	Erythrocyte Sedimentation Rate
ES	Effect Size
EULAR	EUropean League Against Rheumatism
GCP	Good Clinical Practice
GSUS	Grey Scale Ultrasound
HAQ	Health Assessment Questionnaire
HLA	Human Leucocyte Antigen
HR	Hazard Ratio
IL	Interleukin
IQR	InterQuartile Range
JAK	Janus Kinase
JSN	Joint Space Narrowing
MCP	MetaCarpoPhalangeal
MHAQ	Modified Health Assessment Questionnaire

MRI	Magnetic Resonance Imaging
MTP	MetaTarsoPhalangeal
MTX	Methotrexate
NSAIDs	NonSteroidal AntiInflammatory Drugs
NOR-DMARD	NORwegian-Disease Modifying AntiRheumatic Drug Register
OMERACT	Outcome Measures in Rheumatology
OR	Odds Ratio
PDUS	Power Doppler Ultrasound
PGA	Patient Global Assessment
PhGA	Physician Global Assessment
PROMIS	Patient-Reported Outcomes Measurements Information System
RA	Rheumatoid Arthritis
RAID	Rheumatoid Arthritis Impact of Disease
RAMRIS	Rheumatoid Arthritis Magnetic Resonance Imaging Score
RCT	Randomised Controlled Trial
RF	Rheumatoid Factor
PIP	Proximal InterPhalangeal
SD	Standard Deviation
SDAI	Simplified Disease Activity Index
SF-36	Short-Form 36
SF-6D	Short-Form 6 Dimensions
SJC	Swollen Joint Count
SOPs	Standard Operating Procedures
SPSS	Statistical Package for the Social Sciences
SRM	Standardized Response Mean
SSZ	Sulfasalazine
TJC	Tender Joint Count
TNFi	Tumour Necrosis Factor inhibitor
ULRABIT	ULtrasound in Rheumatoid Arthritis patients starting BIologic Treatment
US	Ultrasound
VAS	Visual Analogue Scale
vdHSS	van der Heijde-modified Sharp Score

List of papers

- I. Aga AB, Lie E, Uhlig T, Olsen IC, Wierød A, Kalstad S, Rødevand E, Mikkelsen K, Kvien TK, Haavardsholm EA. Time trends in disease activity, response and remission rates in rheumatoid arthritis during the past decade: results from the NOR-DMARD study 2000-2010. *Ann Rheum Dis* 2015;74(2):381-8.

- II. Haavardsholm EA, Aga AB, Olsen IC, Lillegraven S, Hammer HB, Uhlig T, Fremstad H, Madland TM, Lexberg ÅS, Haukeland H, Rødevand E, Høili C, Stray H, Noraas AL, Hansen IJW, Bakland G, Nordberg LB, van der Heijde D, Kvien TK. Ultrasound in the management of rheumatoid arthritis: results from the randomised controlled ARCTIC trial. [Submitted]

- III. Aga AB, Hammer HB, Olsen IC, Uhlig T, Kvien TK, van der Heijde D, Fremstad H, Madland TM, Lexberg Å, Haukeland H, Rødevand E, Høili C, Stray H, Bendvold AN, Soldal DM, Bakland G, Lie E and Haavardsholm EA. First step in the development of an ultrasound joint inflammation score for rheumatoid arthritis using a data-driven approach. *Ann Rheum Dis* 2015 Published Online First: 17 Jun 2015 doi: 10.1136/annrheumdis-2015-207572.

- IV. Aga AB, Hammer HB, Olsen IC, Uhlig T, Kvien TK, van der Heijde D, Fremstad H, Madland TM, Lexberg Å, Haukeland H, Rødevand E, Høili C, Stray H, Noraas AL, Hansen IJW, Bakland G, Lillegraven S, Lie E and Haavardsholm EA. Development of a feasible and responsive ultrasound joint inflammation score for rheumatoid arthritis through a data-driven approach. [Submitted]

1. Introduction and background

1.1 Rheumatoid arthritis

Rheumatoid arthritis (RA) is a chronic inflammatory disease in which the main feature is joint inflammation with subsequent joint damage and eventually loss of function. The typical clinical picture is symmetric arthritis of the small joints of hands and feet, but larger joints may also be affected. RA can also have extra-articular manifestations and is associated with increased risk of cardiovascular disease and reduced life expectancy (1).

1.1.1 Classification

The former 1987 ACR classification criteria for RA have been criticised for their lack of sensitivity in early disease (2). One of the main rationales for development of the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) classification criteria for RA was to improve sensitivity in early disease (3). Classification as “definitive RA” according to these criteria is based on the confirmed presence of synovitis in at least one joint, absence of an alternative diagnosis better explaining the synovitis, and achievement of a total score of 6 or greater (of a possible 10) from the individual scores in four domains (see table 1). This new classification system focuses on features at earlier stages of the disease that are associated with persistent and/or erosive disease, rather than defining the disease by its late-stage features.

1.1.2 Epidemiology

The prevalence of RA is generally reported to be 0.5–1% in the adult Norwegian population, with an incidence of 25–50 per 100 000 persons per year. The incidence is increasing with age, reaching a plateau around the 5th and 6th decade of life (4-7). RA is more common in women than in men (8). Some data suggest that the incidence of RA has been decreasing, but newer data do not support a continued incidence rate reduction (4, 7, 9). Sixty to eighty percent of patients are positive for rheumatoid factor (RF), an antibody directed against the Fc part of immunoglobulin G (IgG), and a similar proportion have detectable anti-citrullinated peptide antibodies (ACPA) (10). In daily clinical practice ACPA are measured by the anti-CCP test. A meta-analysis published in 2007 found a

sensitivity of 67% and a specificity of 95% for detection of RA by anti-CCP, and a sensitivity of 69% and a specificity of 85% for RA by IgM RF (10).

RA is associated with significant mortality (11), partly caused by a substantially increased risk of cardiovascular disease, in addition to other causes such as gastrointestinal bleeding, lymphoproliferative malignancies and adverse events of RA treatment. However, recent studies indicate that improved treatment strategies reduce the increased mortality risk in early RA patients (12).

Table 1 *The 2010 ACR/EULAR classification criteria for RA* (3)*

JOINTS (0–5) †	
1 large joint	0
2–10 large joints	1
1–3 small joints (large joints not counted)	2
4–10 small joints (large joints not counted)	3
>10 joints (at least one small joint)	5
SEROLOGY (0–3)	
Negative RF <u>AND</u> negative ACPA	0
Low positive RF <u>OR</u> low positive ACPA	2
High positive RF <u>OR</u> high positive ACPA	3
SYMPTOM DURATION (0–1)	
<6 weeks	0
>=6 weeks	1
ACUTE PHASE REACTANTS (0–1)	
Normal CRP <u>AND</u> normal ESR	0
Abnormal CRP <u>OR</u> abnormal ESR	1

* The criteria apply to patients with at least one joint with clinical synovitis not related to a specific aetiology. A patient must achieve a score of 6 or more points to be classified as rheumatoid arthritis.

† Joint involvement includes both tender and swollen joints, and MRI/ultrasound may be used to determine a more complete joint involvement (13).

1.1.3 Pathogenesis

RA is described as a complex genetic disease, meaning that several genes, environmental factors, and chance act in concert to cause pathological events, but the mechanisms are not completely understood (14, 15). Findings in twin studies have estimated the relative contribution of genetic factors to be about 50% for RA, leaving the remaining part to environment and chance (14, 16). The most established environmental risk factor for RA is cigarette smoking (14, 17-20), smokers are also less likely to respond to Methotrexate and TNF-inhibitors (21). It has also been shown that current smoking is a strong predictor of radiographic progression (22). Several studies have documented that there is a dramatic gene-environment interaction between smoking and HLA-DR shared epitope genes in the development of anti-CCP positive RA (23, 24). Other potential environmental factors include silica dust (25), mineral oils (26) and other airway exposures. Infectious agents (e.g. Epstein-Barr virus, cytomegalovirus, proteus species, and Escherichia coli) and their products have long been linked with RA, although underlying mechanisms remain elusive, some form of molecular mimicry is postulated (27, 28). Furthermore, RA appears to be associated with periodontal disease, and specifically the Gram negative bacteria *P. gingivalis* (29).

A wide range of molecules are involved in RA pathobiology, and have become targets for drug development in the last decades (30). The most important are tumour necrosis factor alpha (TNF-alpha), interleukin 6 (IL-6), interleukin 1 (IL-1), janus kinase (JAK), interleukin 17 (IL-17), and spleen tyrosine kinase (SYK) (15, 30). Regulatory authorities have approved drugs targeting the first four for use in RA (JAK inhibitor only approved by the US Food and Drug Administration, FDA and not by European Medicines Agency, EMA). Some of the molecules (e.g. TNF-alpha, IL-6, IL-1) are also directly involved in the development of bone erosions (15). In addition, it is known that T-cell and B-cell activation is involved in RA pathobiology, and a drug that acts as selective T-cell-co stimulator blocker and a cell-depleting agent binding to CD20-antigen on B-lymphocytes are approved.

1.1.4 Treatment

In this section, only pharmacological treatments will be addressed. The main treatment of RA is disease modifying antirheumatic drugs (DMARDs), which are often described as

either synthetic or biologic (31). These terms refer mainly to how the drugs are produced: Biologic DMARDs are created by biologic processes rather than chemical synthesis. Biologic DMARDs have, on a group level, been found both to effectively reduce inflammation in RA and to inhibit joint damage to a greater extent than what could be expected by similarly lowered disease activity with non-biologic DMARDs (14, 32, 33). The different groups of synthetic DMARDs (sDMARDs) and biologic DMARDs (bDMARDs) are presented in table 2, where the first group is divided into conventional synthetic (csDMARDs) and the new targeted sDMARDs (tsDMARDs), and the second group is divided into biologic original DMARDs (boDMARDs) and biosimilars (bbDMARDs) (34). Glucocorticosteroids can be used as supplements to DMARD treatment, administered systemically or as joint injections, and in early RA it is often used as bridging therapy when awaiting the optimal therapeutic effect of sDMARDs (35, 36).

Table 2 Overview of groups of disease-modifying antirheumatic drugs (DMARDs)

Group	Examples
Conventional synthetic DMARDs (csDMARDs)	<i>Methotrexate (MTX)</i> <i>Sulphasalazine (SSZ)</i> <i>Hydroxychloroquine (HCQ)</i> <i>Leflunomide</i> <i>Gold</i>
Targeted synthetic DMARDs (tsDMARDs)	<i>Tofacitinib (Janus Kinase (JAK)-inhibitor)</i> <i>(Not approved by EMA in European Union)</i>
Biologic original DMARDs (boDMARDs)	<i>Adalimumab (TNF inhibitor)</i> <i>Abatacept (selective T-cell co-stimulator blocker)</i> <i>Anakinra (IL-1 inhibitor)</i> <i>Certolizumab pegol (TNF inhibitor)</i> <i>Etanercept (TNF inhibitor)</i> <i>Golimumab (TNF inhibitor)</i> <i>Infliximab (TNF inhibitor)</i> <i>Rituximab (Cell-depleting agent; monoclonal antibody, binds to CD20-antigen on B-lymphocytes)</i> <i>Tocilizumab (IL-6 inhibitor)</i>
Biologic biosimilar DMARDs (bbDMARDs)	<i>SB4, biosimilar etanercept</i> <i>CT-P13, biosimilar infliximab</i>

1.1.5 Treatment strategies

The management of RA has improved greatly over the last decade with the introduction of biologic agents and early initiation synthetic DMARDs (37-46). New treatment strategies have also improved RA care, to an extent where remission has become an achievable goal for a large proportion of RA patients. EULAR and ACR have published treatment recommendations for early RA, last updated in 2013 and 2015, respectively (31, 47). The new treatment strategies are based on the following overarching principles (48-51).

Window of opportunity – *early initiation of DMARDs*

A common sense approach to the management of a persistent, progressive, damaging condition such as RA would seem to intervene before the onset of damage, at a stage when the disease process may still be reversible or at least modifiable. Such a phase of the disease has been described as a “window of opportunity” for intervention (37-39). A growing body of evidence has emphasized the consistent clinical and radiological benefits of early, aggressive treatment of RA (40-42). These and other studies confirm that all therapies – monotherapy, combinations of synthetic DMARD and biologics – work better in early disease than in established RA.

Treat to target – *therapy targeting clinical remission, escalate therapy if target is not reached*

Increasing research has focused on treatment strategies to slow radiographic joint damage progression and prevent loss of function. Several studies have shown that intensive patient management improves disease activity and halts radiographic progression better than routine patient management (table 3) (43-45, 48-50). A paradigm shift was to aim therapeutically for remission, instead of merely improvement in symptoms and signs (43, 46). To reach this goal, treatment is intensified by combining various DMARDs, including biologic agents, and applying a tight control regimen early in the disease (52). MTX is usually selected as initial therapy and as an anchor drug in combination therapies. TNF inhibitors in combination with MTX are usually the first choice of biologic therapy.

Tight control – *frequent assessments of disease activity measures*

Tight control may be defined as a treatment strategy tailored to the disease activity of individual patients with RA with the aim of achieving a predefined level of low disease

activity or preferably remission, within a reasonable period of time (50, 53). Monitoring should be frequent in active disease (every 1–3 months), if there is no improvement by at most 3 months after treatment start or the target has not been reached by 6 months, therapy should be adjusted (46, 47, 54). This strategy with tight control and treat to a pre-defined target mimics the successful approach used for other areas of medicine, such as cardiology and diabetes care and also includes a program for disease monitoring and adjustment of therapy if the pre-defined target is not reached (55).

Table 3: Results from selected treatment strategy trials prior the ARCTIC trial

Study/endpoint	Results (rate)	
	Aggressive strategy	Routine strategy
Fin-RACo (45)		
ACR remission	14%	3%
DAS28 remission	51%	16%
Good treatment response	67%	27%
TICORA (50)		
EULAR remission	65%	16%
Good treatment response	82%	44%
CAMERA (44)		
Clinical remission	50%	37%
CIMESTRA (56)		
ACR remission	35/28%	n/a
DAS28 remission	43/34%	n/a

1.2 Assessment of disease activity and physical function

Assessment of disease activity in RA is crucial in both clinical practice and in research to evaluate the efficacy of treatment, and to monitor the course of the disease (46).

1.2.1 Single outcome measures

Joint counts of tender and swollen joints are necessary to classify an arthritic disease as RA according to both new and old criteria (2, 3), as well as to evaluate the treatment response and achievement of remission in clinical practice and in research (57, 58). The most common formal joint counts are 66 swollen/68 tender joints, the Ritchie Articular Index for tender joints (59) (used in Disease Activity Score [DAS]), 44 swollen joints (used in DAS), and 28 swollen and tender joints (60) (part of the Disease Activity Score with 28 joints [DAS28], Simplified Disease Activity Index [SDAI] and Clinical Disease Activity Index

[CDAI]) (57). The Ritchie Articular Index is a slightly different approach to examination of joint tenderness than the other methods mentioned here, as joints and joint groups are scored 0–3 instead of 0 or 1 (59), and certain joints are treated as a single unit (metacarpal-phalangeal and proximal interphalangeal joints of each hand and the metatarsophalangeal joints of each foot).

Other single outcome measures for assessing disease activity of RA are the patient's assessment of pain, patient's and physician's global assessment of disease activity, the patient's assessment of physical function, and acute phase reactants in the blood (61, 62). When evaluated together, these measures capture the broad range of improvement in RA, and all are at least moderately sensitive to change. Many of them also predict important long-term outcomes in RA, such as physical disability, radiographic damage and mortality (62).

1.2.2 Composite outcome measures

Evaluation of disease activity in RA is not easy, and no single marker can reflect all aspects of the disease. In recent years disease activity instruments, in which single measures are pooled into a composite measure or index, have significantly improved the ability of the rheumatologist to evaluate the course of RA (63). The most commonly used indices are DAS28, DAS, SDAI and CDAI (62, 64-66). These indices are essentially based on the same single measures of disease activity in RA: tender and swollen joint counts, the patient's evaluation of RA activity, acute phase reactants, and, for SDAI and CDAI also the physician's evaluation. Table 4 shows components of each composite disease activity measure.

These indices can be used to monitor disease activity, guide treatment decisions and be used as endpoints in clinical trials. Advantages of a composite measure lies in the increase in responsiveness; possible disadvantages lie in the interpretation of an unfamiliar measure, and in pooling endpoints that measure disparate processes or outcomes (61).

Table 4 Disease activity composite indices in RA

	CRP or ESR	Swollen joint count		Tender joint count		PGA	PhGA
		28 joints	44 joints	28 joints	Ritchie Articular Index		
DAS	x		x		x	x	
DAS28	x	x		x		x	
SDAI	x*	x		x		x	x
CDAI		x		x		x	x
2011 ACR/EULAR Boolean remission criteria	x*	x		x		x	

Abbreviations: ACR: American College of Rheumatology, CDAI: Clinical Disease Activity Index, CRP: C-reactive protein, DAS: Disease Activity Score, DAS28: Disease Activity Score with 28 joints, ESR: Erythrocyte Sedimentation Rate, EULAR: European League Against Rheumatism, PGA: Patient Global Assessment, PhGA: Physician Global Assessment, SDAI: Simplified Disease Activity Index. *CRP.

1.2.3 Remission criteria

The introduction of biologic DMARDs and new treatment strategies in RA has led to an increased focus on RA remission during the last years. Remission, which implies no signs and symptoms of active disease, is now the preferred treatment target in RA, at least in early disease (46). Based on the thought that it is “good to be better, but better to be good”, there is now more often focus on state scores as outcome measures in RA clinical trials, compared to a focus on change scores as in most earlier studies (67). The commonly used RA disease activity states, i.e. remission, low disease activity, moderate disease activity and high disease activity, can be defined based on cut-off values for DAS, DAS28, SDAI and CDAI (63, 68). Table 5 shows thresholds for disease activity states by composite measures.

Table 5 Cut-offs for disease activity states by composite measures for RA

	Disease activity level			
	Remission	Low	Moderate	High
DAS	<1.6	1.6–2.4	>2.4–3.7	>3.7
DAS28	<2.6	2.6–3.2	>3.2–5.1	>5.1
SDAI	≤3.3	>3.3–11	>11–26	>26
CDAI	≤2.8	>2.8–10	>10–22	>22
ACR/EULAR Boolean remission criteria	SJC≤1 TJC≤1 PGA≤1 CRP (mg/L) ≤10	–	–	–

Abbreviations: ACR: American College of Rheumatology, CDAI: Clinical Disease Activity Index, CRP: C-reactive protein, DAS: Disease Activity Score, DAS28: Disease Activity Score with 28 joints, EULAR: European League Against Rheumatism, PGA: Patient Global Assessment, SDAI: Simplified Disease Activity Index.

1.2.4 Biochemical and immunological markers

The ACR core set recommends collection of an acute phase reactant to assess disease activity (62). The most commonly used acute phase reactants are the ESR and CRP. These markers are usually, but not always, elevated in RA patients with active disease. They usually decline with DMARD treatment, and can be used to monitor disease activity over time. Elevated ESR and CRP at the onset of RA are predictive of a more aggressive disease and potentially worse prognosis (69, 70).

RFs are antibodies against the Fc portion of IgG and can be of any immunoglobulin subclass (IgA, IgG and IgM). RF is detectable during the course of disease in about 75–85% of patients with RA, and is predictive of more aggressive erosive disease and poorer long-term function (69, 71).

ACPAs are directed against the citrullinated residues of proteins, and are like RF associated with erosive joint disease and functional disability (69, 72, 73). The anti-CCP test has been shown to be the single most reliable prognostic factor in clinical practice (69, 74).

1.2.5 Physical function and health-related quality of life

Physical function can be assessed by patient reported outcome measures or performance tests (75). Most data collections include one of the versions of the Health Assessment Questionnaire (HAQ). In the original HAQ, also known as the HAQ Disability Index

(HAQ-DI), 20 activities of daily living are queried, with four response categories for each question (76). The answers inform separate values for eight categories, as well as a final single score ranging from 0 (best) to 3 (worst). A shorter modified HAQ (MHAQ), where the number of items is reduced from 20 to 8, is also available (77). A new version of patient reported physical function has recently been developed, the Patient-Reported Outcomes Measurement Information System (PROMIS) physical function, including 20-item short form (78). The original HAQ had 4 response categories, the new PROMIS physical function included a fifth response option, “with a little bit of difficulty”. The total raw score of the 20-item PROMIS physical function is found by the sum of the values of the response to each question (range 20–100). The raw score is translated to a T-score with a mean of 50 and a standard deviation of 10. The physical function components of the SF-36 (79), and the short form-6 dimensions (SF-6D) (80) are other widely used questionnaires measuring functional status in RA. EuroQol-5 Dimensions (EQ-5D) is a standardised instrument for use as a measure of health-related quality of life using three levels of severity in five dimensions with values based on UK preference weights with range from 1 (best possible health), through 0 (death) to –0.59 (worse than death) (81).

1.3 Imaging in rheumatoid arthritis

Imaging modalities capture two main features in RA – structural damage and inflammation (82). Conventional radiography is the oldest radiologic modality and has been in use for a long time to assess joint damage both in RA clinical practice and clinical trials, but the possibility to visualize inflammation on radiographs is very limited. The introduction of modern imaging techniques such as magnetic resonance imaging (MRI) and ultrasound (US) has broadened the imaging modalities in RA from assessment of permanent damage to assessment of ongoing inflammation and all aspects of RA joint pathology (table 6).

Table 6 Conventional radiographs, MRI, Ultrasound and CT, capture of inflammatory and structural features. Table adapted from Haavardsholm et al. (82)

	CR	MRI	US	CT
Inflammatory features				
Synovitis	No direct visualization	Yes – included in RAMRIS (83, 84)	Yes	Yes, not commonly used
Bone marrow oedema/osteitis	No	Yes – included in RAMRIS (83, 84)	No	Yes, not commonly used
Tenosynovitis	No	Yes – scoring system available (85)	Yes	Yes, not commonly used
Structural features				
Erosions	Yes – included in all common RA scoring systems	Yes – included in RAMRIS (83, 84)	Yes (some limitations)	Yes
Joint space narrowing	Yes – included in all common RA scoring systems	Yes – scoring system published by OMERACT (86)	Yes	Yes

Abbreviations: CR: conventional radiograph, MRI: magnetic resonance imaging, US: ultrasound, CT: computer tomography. RAMRIS: Rheumatoid Arthritis Magnetic Resonance Imaging Score

1.3.1 Conventional radiography

Conventional radiographs (CRs) of hands and feet are widely accepted as the gold standard for assessment of structural damage in RA, and assessment of radiographic joint damage is part of the core set of endpoints for RA clinical trials (61). CRs are also recommended for monitoring the disease course in clinical practice (46) and are especially suited to visualize bone structures, including important aspects of RA such as erosions and joint space narrowing. Radiographs are obtained and evaluated by radiologists in routine clinical care, but formal scoring of pathologies is usually only applied in research studies.

Scoring systems

John Sharp et al. proposed one of the first scoring methods for RA in 1971, revised in 1985 (87, 88). The scoring method assessed erosions and joint space narrowing, including ankylosis, in the hand and wrist (88). It has later been modified several times (87). The most

well-known modifications are the Genant and the van der Heijde Sharp modifications, and these are the predominant choices in RA research today (89). The van der Heijde-modified Sharp score (vdHSS) include both hands and feet, and score erosions and joint space narrowing separately. The maximum scores for the vdHSS are outlined in table 7 (90). There are several other scoring systems for conventional radiographs, i.e. the Larsen score and Ratingen score (87, 91).

Table 7 Maximum scores according to the van der Heijde-modified Sharp score (90)

	Hands	Feet	Total
Erosion	160	120	280
Joint space narrowing	120	48	168
Total score	280	168	448

1.3.2 Ultrasound

Musculoskeletal ultrasound (US) is increasingly being used to detect and monitor joint inflammation in RA, both in clinical practice and in research. Some of the advantages of US are the relatively low cost, good accessibility, patient-friendliness and possibility to assess multiple joints at point of care in a relatively short time period. The examination involves no ionizing radiation and does not involve intravenous contrast.

Structured ultrasound examination has been a promising tool for monitoring of disease activity in RA (92-94). US can assess two aspects of synovitis: Morphology and quantity by grey-scale (GSUS) and synovial vascularity by power Doppler (PDUS). Other aspects of pathology that potentially can be visualized using US include bone erosions, cartilage, osteophytes, tenosynovitis and enthesitis (95-99). Musculoskeletal ultrasound pathologies have been defined by the OMERACT 7 Special Interest Group (100). The ability to perform “dynamic” evaluation of joints and tendons and help in guiding needle positioning in interventional procedures are some of the other reasons for the rapid implementation of US in the management of RA.

Scoring systems

Compared to the situation within conventional radiography and magnetic resonance imaging, there has been a lack of consensus on a scoring system for ultrasound, especially

there is no agreement on the optimal number and which joints and tendons should be assessed for a sensitive and feasible US assessment of joint inflammation (93). Szkudlarek et al. have published a 0–3 semiquantitative scoring system for joint effusion, synovial thickening and bone change for finger and toe joints which is widely used (101). A validated 0–3 semiquantitative scoring system for GSUS and PDUS for 32 joints was published in 2011, with standardised probe placements and patient position and with an US atlas as reference (102). This scoring system has shown high inter- and intra-reader reliability, and is used in most of the rheumatology departments in Norway and also in international studies (103, 104).

A complete US examination of all joints and tendons that can be affected in RA would be extremely time consuming and therefore infeasible. To increase feasibility, several reduced joint scores for assessing joint inflammation have been developed (94, 105-112). The German 7-joint ultrasound score of the dominant hand and foot developed by Backhaus et al. is one of the best known reduced scores. Wrist, MCP2, MCP3, PIP2, PIP3, MTP2, MTP5 and associated tendons are evaluated by a semiquantitative system for GSUS (0–3), PDUS, (0–3) erosions (0–1) and tenosynovitis (0–1). Most of the joints and tendons are examined in several projections (105). Naredo et al. developed a 12-joint score by a data reduction procedure from 44 joints based on the frequency of joint involvement at baseline. Bilateral examinations include the following 6 joints: elbow, wrist, MCP2, MCP3, knee and ankle. Each joint is scored for synovitis and PDUS signal on a scale from 0 to 3 (108). The 6-joint score developed by Perricone et al. is a further reduction of the previous mentioned 12-joint score. The joints included are bilaterally MCP2, knee and ankle, and the selection was based on the frequency of synovial site involvement by synovial effusion, synovial proliferation and PDUS (109). Some of these scoring systems have shown good agreement with more comprehensive ultrasound joint scores, but to our knowledge they were not developed by purely data-driven approaches (108, 113).

Further evaluation of specific joints to be assessed is included in the research agenda proposed in the recent EULAR recommendations for the use of joint imaging in the clinical management of RA (114).

1.3.3 Magnetic resonance imaging

Magnetic resonance Imaging (MRI) is a non-invasive tomographic imaging technique that yields cross-sectional images in any plane. MRI allows simultaneous examination of all components of the joint, including synovial membrane, intra- and extra-articular fluid collections, soft tissues, articular cartilage, ligaments, tendons, tendon sheets and bone, without ionization radiation and with minimal adverse effects (82, 115). MRI visualizes both joint destruction and joint inflammation. MRI is also the only imaging technique that can visualize bone marrow oedema (also referred to as osteitis), one of the strongest imaging predictors for future joint damage (116-119).

Scoring systems

The OMERACT (Outcome Measures in Rheumatology) RA MRI scoring system (RAMRIS) is validated for assessment of MRI images of the wrist and MCP joints in RA (83, 120). RAMRIS includes a core set of MRI sequences and a semiquantitative scoring system for synovitis, bone erosions and bone marrow oedema (83), and a reference atlas has been developed to aid the scoring process (121). More recently, a joint space narrowing score was also validated (122). A scoring system for tenosynovitis has also been published, which is used as an addendum to the OMERACT scoring system (84). The RAMRIS has been proven valid, reliable and sensitive to change (84, 123) and is increasingly used as an outcome measure in clinical trials (124, 125).

1.3.4 Imaging remission in rheumatoid arthritis

Clinical remission is now a defined treatment target in current RA treatment recommendations (47, 54). The introduction of biologic therapies in combination with aggressive treatment strategies has led to levels of RA disease control that were previously not achievable (126-132), and remission has become a realistic goal of therapy (46, 47, 54, 133-136). However, studies have shown that progression of radiographic joint damage may occur in clinical remission regardless of choice of remission definition. Subclinical inflammation detected by modern imaging techniques such as ultrasound and MRI is present in the majority of patients in clinical remission, and is associated with progressive joint damage and disease activity flare in these patients.

In a study by Brown et al. 102 patients in clinical remission were assessed by various imaging modalities. Despite being in clinical remission (according to clinical judgement), 19% of patients deteriorated radiographically, and this progression was largely explained by PDUS signal (with a 12 times higher odds ratio for progression in joints with increased PDUS signal) (137). In a study by Scire et al. of 106 early RA patients in clinical remission followed for 24 months, positive PDUS signal, even in a single joint, was the main predictor of relapse within 6 months (multivariate logistic regression analysis) (138).

Bone marrow oedema visualised by MRI is one of the strongest imaging predictors of future joint damage (116-119, 139). In a study by Gandjbakhch et al, 85 RA patients in remission or low disease activity were followed for one year. Bone marrow oedema at baseline was predictive of structural progression on MRI (140). This finding confirmed the prognostic role of bone marrow oedema with regard to subsequent bone erosions, even in RA patients in remission or low disease activity (140).

In a multi-centre study coordinated by the OMERACT MRI in Inflammatory Arthritis group data from five cohorts were collected (141). RA patients in clinical remission (N=185) or low disease activity (N=69) with available MRI and conventional radiograph scores were included. The authors found that high MRI synovitis score predicted radiographic progression in patients in clinical remission/low disease activity, and they established a cut-off point for determining an MRI inflammation activity acceptable state based on the RAMRIS synovitis score of RF-positive RA patients.

In summary, remission classified according to established clinical criteria is not an inflammation-free, non-damaging disease state. Østergaard and Møller-Bisgaard addressed the question: Is imaging needed to define remission in rheumatoid arthritis? (142) No studies have yet documented that an imaging-guided treatment strategy will improve patient outcome over and above a treatment strategy based on conventional tight control regimen. However, in addition to the ARCTIC trial included in this thesis, we are aware of one other trial (NCT00920478) designed to assess the value of ultrasound in RA, and two other ongoing randomised controlled trials of imaging-guided treatment strategies in RA (table 8) (142).

Table 8 Overview of clinical trials of imaging-guided treatment strategies for rheumatoid arthritis Table adapted from Østergaard et al. (142)

Study	Status (Location)	ClinicalTrials.gov identifier	Target	Primary endpoint(s)
Targeting Synovitis in Early Rheumatoid Arthritis (TaSER)*	Completed (Scotland)	NCT00920478	Ultrasound-based (PDUS signal in ≤ 1 joint)	MRI RAMRIS erosion score and DAS28
Aiming for Remission in Rheumatoid Arthritis (The ARCTIC trial)†	Completed (Norway)	NCT01205854	Ultrasound-based (no PDUS signal)	Complete clinical remission (i.e. DAS<1.6, no swollen joints and no radiographic progression during the 8 last months)
An MRI-guided Treatment Strategy to Prevent Disease Progression in Patients With Rheumatoid Arthritis (IMAGINE-RA) ‡	Ongoing, not recruiting (Denmark)	NCT01656278	MRI-based (no bone marrow oedema)	Clinical remission and no radiographic progression
Targeted Ultrasound in Rheumatoid Arthritis (TURA) §	Ongoing, recruiting (Europe)	NCT02056184	PDUS-based (no PDUS signal)	PDUS (decrease in PDUS signal)

Abbreviations: DAS: 44-joint disease activity score, PDUS: power Doppler ultrasound, RAMRIS: rheumatoid arthritis MRI scoring system.

* US National Library of Medicine. *ClinicalTrials.gov* [online], <http://clinicaltrials.gov/ct2/show/NCT00920478> (2009).

† US National Library of Medicine. *ClinicalTrials.gov* [online], <http://clinicaltrials.gov/ct2/show/NCT01205854> (2013).

‡ US National Library of Medicine. *ClinicalTrials.gov* [online], <http://clinicaltrials.gov/ct2/show/NCT01656278> (2014).

§ US National Library of Medicine. *ClinicalTrials.gov* [online], <http://clinicaltrials.gov/ct2/show/NCT02056184> (2014).

2. General aim and research questions

2.1 General aim

The general aim of this thesis was to assess the implementation of modern treatment strategies for RA in clinical practice during the last decade, focusing on the change towards tight control, treating disease activity to target and the value of ultrasound in clinical decision-making in early RA.

2.2 Research questions

- How did baseline disease activity levels and disease duration in patients with RA starting MTX monotherapy and first TNFi+MTX therapy, respectively, change from 2000 to 2010? (Paper I)
- How did response and remission rates in patients with RA starting MTX monotherapy and first TNFi+MTX therapy, respectively, change from 2000 to 2010? (Paper I)
- What were the remission rates after two years of DMARD treatment in patients with early RA treated according to current recommendations, including tight control and treat to target, in the time period 2010–2015? (Paper II)
- Is ultrasound imaging remission feasible in early RA? (Paper II)
- Was there any added effect of applying ultrasound versus not applying ultrasound in a clinical tight control regimen in patients with early RA with respect to achievement of sustained clinical remission? (Paper II)
- Was there any added effect of applying ultrasound versus not applying ultrasound in a clinical tight control regimen in patients with early RA with respect to achievement of radiographic no-progression? (Paper II)
- How much of the total information was retained by a reduced ultrasound inflammation scores in a separate cohort of established RA? (Paper III)

- How was the responsiveness of a novel ultrasound joint inflammation score? (Paper IV)
- How did a novel ultrasound joint inflammation score perform regarding sensitivity to change and total information retained compared to formerly proposed ultrasound scores in a separate validation cohort of RA? (Paper III and IV)

3. Material and methods

3.1 Study design

Randomised controlled trials (RCTs) and meta-analyses of RCTs are classified as the study types that provide the highest category of evidence (143-145). However, strict inclusion and exclusion criteria and short follow-up time limit the external validity of these studies (146). Data from longitudinal observational studies (LOS) are important complements to the results from RCTs, especially regarding effectiveness and information about safety in clinical practice.

Data from LOS were used for the analyses in the first, third and fourth paper, and data from a RCT were used for the analyses in the second, third and fourth paper included in this thesis. The next sections describe the study populations and organization of data collection of each cohort. In this thesis patients from three study populations were included: the NOR-DMARD study, the ULRABIT study and the ARCTIC trial.

3.2 Study populations and organization of data collection

3.2.1 The NOR-DMARD study

The NOR-DMARD study is a longitudinal multi-centre observational study, where adult patients (>18 years of age) with inflammatory joint diseases starting a treatment regimen with synthetic DMARDs or biologic DMARDs are included. Patients are followed only during treatment, and re-included if switched to another treatment. Thus, each case represents one treatment regimen. In this current thesis we included RA patients from

NOR-DMARD starting MTX monotherapy or first biologic DMARD in the time period 2000 to 2010, with two year follow-up data.

The five centres recruiting patients in this time-period were Diakonhjemmet Hospital, Lillehammer Hospital for Rheumatic Diseases, University Hospital of North Norway in Tromsø, Drammen Hospital Vestre Viken HF and St Olavs Hospital in Trondheim, covering more than 1.5 million inhabitants, nearly thirty percent of the Norwegian population. The completeness of the register has been approximately 85%; the remaining 15% were either missed for inclusion, refused enrolment or were excluded due to language barriers or inclusion in ongoing randomised controlled trials. A total of 10,876 treatment courses in more than 7,300 individual patients had been included in the NOR-DMARD study per December 2010. Patients are still recruited to the NOR-DMARD study, but with a new protocol from 2012. The study is registered in the ClinicalTrials.gov database (NCT01581294). Rheumatology care in Norway is largely hospital based, making such a design good for completeness and representativeness of included patients. Access to biologic therapies is good, and fully financed by the public health care system.

Organisation of data collection

The data were collected at baseline, 3, 6 and 12 months, and annually thereafter. In most centres patients filled in all patient-reported measures and were seen by a study nurse who performed joint-counts and recorded all required study information, before they were seen by their physician. The study nurses checked the completeness of the registrations. Data management and some monitoring were performed by a contract research organization (Smerud Medical Research), and researchers were provided with three database updates per year. Further validation and data management was subsequently conducted by researchers utilising the data.

3.2.2 The ULRABIT study

The ULRABIT study (ULtrasound in Rheumatoid Arthritis patients starting BIologic Treatment) is a single-centre longitudinal observational study at Diakonhjemmet Hospital, and a supplemental study to NOR-DMARD. Patients with established RA with active disease starting or switching a new biologic DMARD treatment were included between 2010 and 2012 in both ULRABIT and NOR-DMARD, and followed for 12 months. Two-

hundred and twelve patients were included. The study is registered in the Anzctr.org.au database (ACTRN12610000284066).

Organisation of data collection

Patients were examined at baseline, 1, 2, 3, 6 and 12 months with a comprehensive ultrasound joint/tendon examination performed by an experienced sonographer (Hilde Berner Hammer) after a validated semiquantitative scoring system of 36 joints and 4 tendons (102), and an experienced study nurse (Anne Katrine Kongtorp) performed joint-counts at every visit.

Data management has been done by a statistician (Inge Christoffer Olsen) and the researchers.

3.2.3 The ARCTIC trial

The ARCTIC trial (Aiming for Remission in rheumatoid arthritis: a randomised trial examining the benefit of ultrasound in a Clinical Tight Control regimen) was a 24-month randomised, open, parallel group clinical strategy study conducted in 11 centres in Norway. The study was designed to compare two tight control treatment strategies for early RA to assess whether incorporation of ultrasound information into treatment decisions, as well as targeting therapy towards imaging remission, would lead to improvement in a combined outcome of sustained clinical remission, absence of swollen joints and inhibition of joint damage. The study is registered in the ClinicalTrials.gov database (NCT01205854).

The centres that recruited patients were four rheumatology departments at University hospitals, six regional/community hospitals and one private practice: Diakonhjemmet Hospital, Haukeland University Hospital, St Olavs Hospital in Trondheim, University Hospital of North Norway in Tromsø, Ålesund Hospital, Drammen Hospital Vestre Viken HF, Martina Hansens Hospital, Hospital Østfold HF in Moss, Haugesund Rheumatism Hospital AS, Sørlandet Hospital HF, and The Rheumatology Clinic Dovland/Bendvold.

Primary endpoint

The primary endpoint was complete clinical remission at the end of the study, and was defined as the proportion of patients meeting all of the following three criteria:

- Disease Activity Score <1.6 at 16, 20 and 24 months

- Absence of swollen joints at 16, 20 and 24 months
- No radiographic progression (<0.5 units change in van der Heijde-modified Sharp Score) between 16 months and 24 months

Inclusion and exclusion criteria

Inclusion and exclusion criteria are outlined in table 9.

Table 9 Main inclusion and exclusion criteria in the ARCTIC trial

Inclusion criteria Patients must meet all of the following:	Exclusion criteria Patients with any of the following would not be eligible:
Male or non-pregnant, non-nursing female	Abnormal renal function
>18 years and <75 years of age	Abnormal liver function
Patients classified as having RA (according to 2010 ACR/EULAR criteria) (3)	Major co-morbidities as outlined in the protocol
Disease duration less than 2 years (defined as time from 1 st patient reported joint swelling)	Leukopenia and/or thrombocytopenia
The treating rheumatologist decides the patient requires DMARD-treatment	Inadequate birth control conception, pregnancy, and/or breastfeeding
The patient has taken no prior DMARD	Indications of active tuberculosis
Patients able and willing to give written informed consent and comply with the requirements of the study protocol	Psychiatric or mental disorders, alcohol abuse, other substance abuse, language barriers, other factors making adherence to the study protocol impossible

Abbreviations: ACR: American College of Rheumatology, DMARD: disease modifying anti-rheumatic drug, EULAR: European League Against Rheumatism, RA: rheumatoid arthritis

Organisation of data collection

Patients were assessed at baseline, 1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 20 and 24 months. Patients recorded patient-reported measures, joint counts were performed by experienced study nurses or physicians, and other relevant examinations and required study information were assessed by physician or study nurse as appropriate. The patients in the ultrasound strategy

arm were assessed by ultrasound at every visit by experienced sonographers, according to a validated scoring system of 32 joints (102). Patients in the conventional tight control arm were assessed by ultrasound yearly, but both the patient and treating physician were blinded to the results.

Treatment protocol and targets

The ultrasound tight control strategy targeted ultrasound imaging remission, defined as no ultrasound power Doppler signal in any assessed joint, in combination with clinical remission, while the clinical tight control strategy was targeted towards clinical remission only. Clinical remission was defined as Disease Activity Score (DAS) <1.6 and no swollen joints. The treatment in both arms was escalated according to the same algorithm (table 9). The initial therapy was methotrexate 15 mg/week increased to 20 mg/week by week five, in combination with 7 weeks of prednisolone with tapering doses from 15 mg to zero. Further steps in the treatment algorithm included methotrexate 25 mg/week, triple synthetic DMARD therapy (methotrexate, sulphasalazine, hydroxychloroquine) and biologic treatment according to guidelines (table 10). In both arms, swollen joints were treated by intra-articular steroids, in addition to any joint with power Doppler signal in the ultrasound tight control arm. All injections in the ultrasound tight control arm were guided by ultrasound, while ultrasound guided injections were not allowed in the conventional arm.

The clinical decision to adjust medication was based on level and change in DAS. In the ultrasound arm the physician should overrule the DAS based decision and proceed to the next treatment regimen if indicated by the ultrasound score. If the patient did not respond adequately (table 10), the physician immediately adjusted the therapy by proceeding to the next step in the treatment algorithm. Current medication was continued if the treatment response was satisfactory. Study personnel entered information on DAS values at the present and previous visit, as well as the ultrasound score at both visits if applicable, into a web page that automatically applied the treatment decisions rules in the appropriate study arm (table 11), and the treating physician was presented with the outcome of the response assessment and whether to escalate treatment or not.

Table 10 *Treatment protocol in the ARCTIC trial*

Visit (months)	Treatment if no response (if response continue treatment at present step, see table 11)
1 (0)	A. Monotherapy* + Prednisolone: 1. Methotrexate 15 mg/week, increase by 2.5 mg every 2 nd 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) 2. Concomitant folic acid 5 mg/week (1mg 5/7 days or 5 mg x 1/week) 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 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)* Joint injections allowed as indicated according to treatment arm.
3 (2)	A. Optimize monotherapy* Increase Methotrexate to 25–30 mg/week <i>Or increase sulphasalazine/hydroxychloroquine/leflunomide dose</i>
4 (3)	A. Monitor start-up regimen (no changes in medication allowed unless due to AE)* Joint injections allowed as indicated according to treatment arm.
5 (4)	B. Triple combination therapy (or other combination therapy if MTX not tolerated):† 1. Add sulphasalazine, step up over 4 weeks to 500mg 2 x 2 and 2. Add hydroxychloroquine 200mg 1 x 2
6 (6)	B. Optimize triple combination therapy:‡ Add Prednisolone 7.5 mg 1 x 1
7 (8)	C. DMARD‡ and 1st biologic:§ 1. Highest tolerable dose MTX* and 2. Add 1 st biologic (according to current Norwegian recommendations) <i>*Or sulphasalazine/hydroxychloroquine/leflunomide if MTX not tolerated</i>
8 (10)	C. DMARD and 1st biologic: Adjust dose/interval of 1 st biologic
9 (12)	D. DMARD‡ and 2nd biologic: Switch to 2 nd biologic (according to current Norwegian recommendations)
10 (14)	D. DMARD‡ and 2nd biologic: Adjust dose/interval of 2 nd biologic
11 (16)	E. DMARD‡ and 3rd biologic: Switch to 3 rd biologic (according to current Norwegian recommendations)
12 (20)	E. Optimize DMARD and 3rd biologic plus prednisolone: Adjust dose/interval of 3 rd biologic and/or add prednisolone 7.5mg
13 (24)	F. Continue medication according to standard clinical care

* If MTX is not tolerated, switch to subcutaneous methotrexate), then continue according to scheme. In case of AE or not tolerated even in low dose subcutaneous, switch to sulphasalazine or hydroxychloroquine 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 oedema 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 synthetic DMARD, this can be omitted if the biologic DMARD chosen has indication for monotherapy (e.g. tocilizumab).

§ 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).

Table 11 ARCTIC decision rules*

	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	
Ultrasound 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

* To be applied at all visits except visit 2 and visit 4.

† 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 hexacetonide). In the ultrasound tight control group joints with PD-signal on US is an additional target.

Data management and study documentation

Data were collected and organized by a web-based electronic case report form (eCRF), (VieDoc™, Pharma Consulting Group, Uppsala, Sweden). Data management and statistical analyses of the primary outcome have been performed by statistician Inge Christoffer Olsen (in collaboration with the other researchers).

Table 12 Comparison of the NOR-DMARD study, ULRABIT study and ARCTIC trial

	NOR-DMARD	ULRABIT	ARCTIC
Study design	Longitudinal observational study	Longitudinal observational study	Randomised controlled trial
Years of inclusion	2000–2010	2010–2013	2010–2013
Number of patients included	In total: >10 000 RA: 6218 MTX naïve, MTX monotherapy: 1866 Biologic naïve, TNFi+MTX: 707	212	238
RA diagnosis	Physician diagnosis (ICD-10 codes M05.8, M05.9 and M06.0)	Physician diagnosis (ICD-10 codes M05.8, M05.9 and M06.0)	Fulfilment of ACR/EULAR 2010 classification criteria
Timing of visits	Baseline, 3, 6, 12 months, thereafter yearly	Baseline, 1, 2, 3, 6 and 12 months	Baseline, 1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 20, 24 months
Imaging protocol	No	Ultrasound	Ultrasound Conventional radiographs MRI
Treatment	According to treating physician's preference and local and national recommendations	Biologic DMARD according to treating physician's preference and national recommendations	Predefined treatment protocol with tight control and treat to target, described in 3.2.3.

Table 13 Use of cohorts in the papers of this thesis

	Paper I	Paper II	Paper III	Paper IV
NOR-DMARD	x			
ULRABIT			x	x
ARCTIC		x	x	x

3.3 Data collection

3.3.1 Demographic and other background variables

The data collection in NOR-DMARD, ULRABIT and ARCTIC included demographic variables such as age, gender, co-morbidities, current medication, disease duration, smoking status, alcohol consumption, education, work status and marital status. Height and weight were measured in ARCTIC, and BMI was calculated. Immunological markers (anti-CCP and RF) were measured in all cohorts.

3.3.2 Clinical, serological and physical outcome measures

A number of outcome measures were recorded at baseline and follow-up. The outcome measures and whether they were utilized in the papers included in this thesis are shown in table 14.

Table 14 *Included assessments and overview of use in included articles*

Collected outcome measures	Paper I	Paper II	Paper III	Paper IV
Joint counts*	x	x	x	x
ESR and CRP	x	x	x	x
Investigator's global assessment of disease activity†	x	x	x	x
Patient's global assessment of disease activity†	x	x	x	x
Patient pain assessment†	x	x		
Patient fatigue assessment†	x	x		
MHAQ	x			
PROMIS physical function		x		
SF-36		x		
SF-6D	x			
EQ-5D		x		
Composite outcome measures				
DAS28	x		x	x
DAS		x	x	x
SDAI	x	x		
CDAI	x			
DAS28 remission (<2.6)	x			
DAS remission (<1.6)		x		
SDAI remission (≤3.3)	x	x		
CDAI remission (≤2.8)	x			
ACR EULAR remission	x	x		
EULAR response criteria	x	x		
Imaging				
Ultrasound of 36 joints and 4 tendons		x	x	x
Conventional radiographs of hands and feet		x		

* 28 tender and swollen joint counts + ankles and forefeet (metatarsophalangeal joints assessed as a single unit bilaterally) in the NOR-DMARD and the ULRABIT study. 44 swollen joint count and Ritchie articular index were assessed in the ARCTIC trial.

† Scored on a 0–100 mm VAS scale

Abbreviations: ESR: Erythrocyte Sedimentation Rate, CRP: C-reactive protein, MHAQ: Health Assessment Questionnaire, PROMIS: Patient-Reported Outcomes Measurements information, SF-36: short form 36, SF-6D: short form 6 dimensions, EQ-5D: EuroQol-5 dimensions, DAS: disease activity score, CDAI: clinical disease activity index, SDAI: simplified clinical disease activity index, ACR: American College of Rheumatology, EULAR: European League Against Rheumatism.

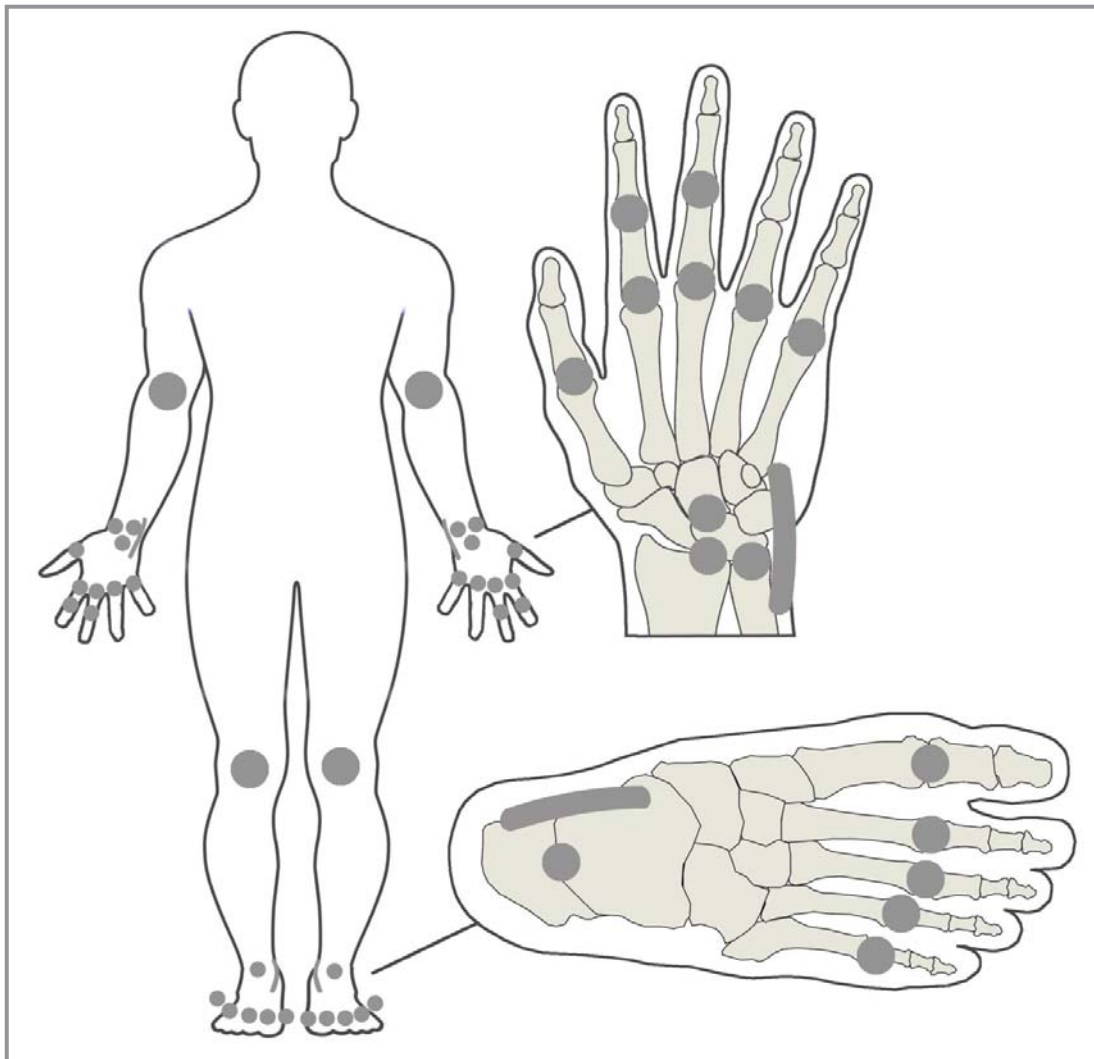
3.3.3 Ultrasound

An extensive ultrasound examination was performed by experienced sonographers using 0–3 semiquantitative scoring systems for both GSUS and PDUS in each of the following 36 joints and 4 tendons: metacarpophalangeal (MCP) 1–5, proximal interphalangeal (PIP) 2–3, radiocarpal, intercarpal, distal radioulnar, elbow, knee, talocrural, metatarsophalangeal (MTP) 1–5, extensor carpi ulnaris (ECU) tendon and tibialis posterior (TP) tendon bilaterally (Figure 1) (102, 147, 148). The range of the sum scores was 0–120 for both GSUS and PDUS and the scanning protocol was a slight modification of a previously published 32-joint protocol (with addition of bilateral PIP2–3, ECU and TP tendons) with the same probe placement and patient positioning and with an ultrasound atlas as a reference (Figure 2) (102).

The ultrasound examinations in the ARCTIC trial were performed by several sonographers (multi-centre study), while a single examiner (Hilde Berner Hammer) performed all examinations in the ULRABIT-cohort (single-centre study). All the sonographers in the multi-centre study underwent training in the form of an ultrasound workshop with both static and dynamic hands-on exercises to calibrate readers, and the workshop was repeated yearly (102). This validation study showed high inter-observer and intra-observer reliability, and most of the examiners in this study were also examiners in our multi-centre study.

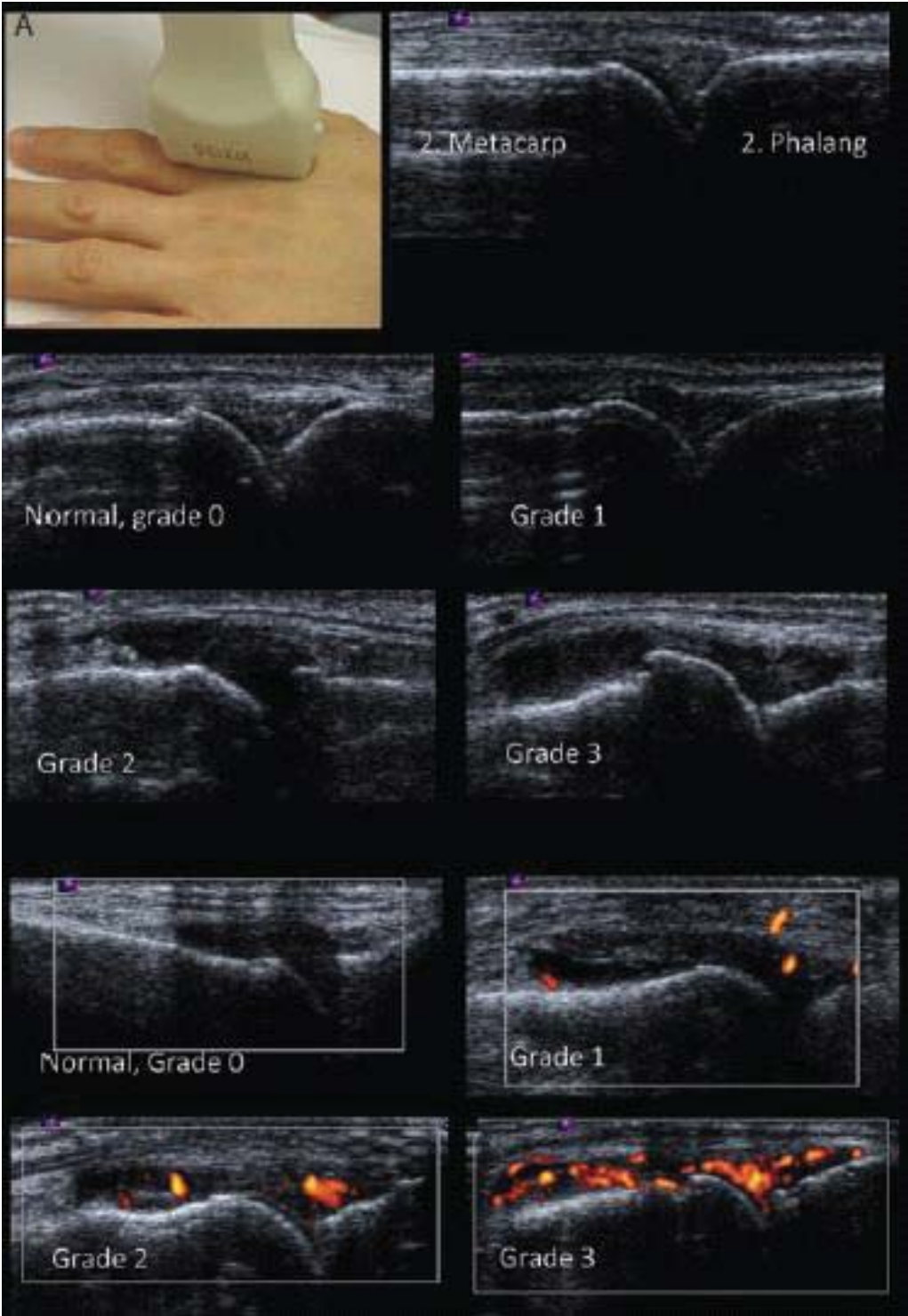
Siemens Antares Sonoline machines (Siemens Medical solutions, Mountain view, CA, USA) with linear probes (5–13 MHz and setting at 11.4 MHz) and identical settings optimized for PDUS in superficial joints (pulse repetition frequency 391 Hz, low wall filter and frequency 7.3 MHz), or GE Logiq E9 (GE Medical Systems Ultrasound and Primary Care Diagnostics, Wauwatosa, WI, USA) with linear probes (setting at 13 MHz) and identical settings for PDUS in superficial joints (pulse repetition frequency 600, power Doppler frequency 10.0 MHz), were used in all the 11 hospitals for the ultrasonography assessments (149). The ultrasound machines were calibrated and optimized for power Doppler sensitivity using technical personnel from the manufactures, to ensure correct settings in all machines.

Figure 1 *Ultrasound joint and tendon examination*



36 joints and 4 tendons examined by GSUS and PDUS: metacarpophalangeal (MCP) 1–5, proximal interphalangeal (PIP) 2–3, radiocarpal, intercarpal, distal radioulnar, elbow, knee, talocrural, metatarsophalangeal (MTP) 1–5, extensor carpi ulnaris (ECU) tendon and tibialis posterior (TP) tendon bilaterally.

Figure 2: Excerpt from the reference atlas used in the ARCTIC and ULRABIT studies



MCP2 joint, probe position and grade 0–3 for GSUS and PDUS. Images courtesy of Dr. Hilde Berner Hammer.

3.3.4 Conventional radiographs

Radiographs of hands, wrists, and feet were scored by two trained readers affiliated to Imaging Rheumatology in the Netherlands, blinded for clinical data and treatment strategy, and radiographs were read independently in chronological order according to the van der Heijde-modified Sharp score (subscores for erosions (0–280) and joint-space narrowing (0 to 168), total score range 0 to 448, higher scores indicating more joint damage) (90).

3.3.5 Adverse events and discontinuations

In NOR-DMARD and ULRABIT date and cause of discontinuation of the current DMARD-regimen was recorded. Causes include «loss/lack of efficacy» (LOE), «adverse events» (AE), «combination of LOE and AE», «patient preference», «remission» and «other/unknown (specified)». Two-year retention to therapy (drug survival) data were analysed and reported in Paper I.

In ARCTIC safety was evaluated by assessment of clinical and laboratory adverse events and serious adverse events were coded according to the Medical Dictionary for Regulatory Activities (MedDRA, v.17.0). Adverse event data were analysed and reported in Paper II.

3.4 Statistics

All statistical analyses in paper I were performed using the Predictive Analytics Software program, V.19 (SPSS, Chicago, Illinois, USA), in Paper II the statistical analyses were performed in Stata v14 (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX, USA). In paper III and IV the factor analyses and regression analyses according to maximising the R^2 were performed using SAS software package (V.9.2 SAS institute, Cary, North Carolina, USA), and the rest of the statistical analyses were performed using SPSS (V.21, SPSS, Chicago, Illinois, USA).

All analyses in this thesis have been performed in collaboration with a statistician.

3.4.1 Descriptive statistics

Continuous measures that were found to be close to normally distributed are presented as means with standard deviation (SD). Skewed measures are presented as medians with the

25th percentile – 75th percentile (sometimes referred to as the interquartile range [IQR]). Dichotomous variables are primarily presented as percentages.

3.4.2 Group comparison

Group comparisons of continuous measures that were close to normally distributed were performed by independent t-tests. Skewed measures were compared using Mann-Whitney U-test. Dichotomous measures were compared by Chi²-test.

3.4.3 Multivariate analyses

In paper I time trends in several baseline variables were assessed by linear regression analysis with year at onset of treatment as independent variable (continuous 1–10) and the respective baseline variables as dependent variables.

In paper II logistic regression analysis was used to assess the primary and other categorical endpoints. Estimates of risk difference in paper II were calculated from the logistic regression parameters using the delta method to provide confidence intervals. We used median regression with 10 000 bootstrap replications to estimate the difference in median radiographic change scores. To analyse other continuous change variables, we used analysis of covariance adjusted for baseline value in addition to study centre and anti-CCP. Estimates of treatment difference were calculated using least-square means.

In paper III and IV we used linear regression analysis to estimate the proportion of information retained (R^2) by the reduced ultrasound scores from the total ultrasound score, with the total ultrasound score as dependent variable, and the simple sum ultrasound score of the candidate sets as the independent variable.

3.4.4 Survival analysis

In paper I drug survival was assessed by Kaplan-Meier analysis. Analyses were limited to the first 2 years after start of treatment to avoid making assumptions on a very limited number of patients. Group comparisons were performed by log-rank test.

3.4.5 Factor analysis

Factor analysis is a variable reduction procedure. It is useful when you have obtained data of a number of variables and believe that there is redundancy among those variables, and

that some of the variables are correlated with each other, often because they are measuring the same construct. We used factor analyses in paper III and IV, to identify joints with high within-group correlation and low between-group correlation. The factor analyses were based on the principal component methodology, with the factors rotated according to the varimax criteria. The number of factors was based on eigenvalues with a cut-off of one. Based on the factor analyses on GSUS and PDUS, we identified groups with correlating scores, and based on these results we formed candidate sets of joints/tendons.

3.4.6 Responsiveness

The term responsiveness denotes the magnitude of change or sensitivity to change over time. We assessed responsiveness using standardised response means (SRMs) in paper III and IV, as there is evidence for this being a valid method, as well as it is the most widely used responsiveness statistic (150-152). There is no universal agreement on how to interpret the magnitude of the SRMs, but in most cases the thresholds introduced by Cohen for effect sizes (ES) are applied: "trivial" ($ES < 0.20$), "small" ($ES \geq 0.20$ and < 0.50), "moderate" ($ES \geq 0.50$ and < 0.80), or "large" ($ES \geq 0.80$) (153).

The SRMs were calculated as the mean change divided by the standard deviation of the change, and their 95% confidence intervals were calculated by applying bootstrapping with 5000 replications.

3.4.7 Cumulative probability plot

A cumulative probability plot depicts data from each individual patient ranked from lowest to highest observed score within each subgroup. In our case, change in radiographic scores was indicated on the y-axis and the cumulative percentage of patients on the x-axis. This is a preferred method to depict radiographic progression in different patient groups (154), and was used in paper II.

3.4.8 Sample size estimation in the ARCTIC trial

The sample size was estimated based on the assumption that the rate of reaching the primary endpoint was 45% in the non-ultrasonography group and 65% in the ultrasonography group, resulting in a treatment difference of 20%. Based on results from previous strategy studies in RA (presented in table 3), a total sample size (study completers) of 198 (99 in each

group) was needed to achieve 80% power to detect a difference of 20% between the groups in a two-sided test at 5% significance level. We aimed to include 240 patients to compensate for attrition. Efficacy and safety analyses included data from all randomised patients who initiated the allocated intervention by attending at least one regular visit after randomisation (the full analyses set, see section 3.4.10).

3.4.9 Hypotheses and decision rules in the ARCTIC trial

Statistical hypothesis

The ARCTIC protocol was designed to establish the superiority of applying ultrasound to not applying ultrasound in a clinical tight control regimen in patients with early RA for the primary endpoint after 24 months of treatment. The null hypothesis was that there is no difference in the probability of achieving the primary endpoint after 24 months of treatment between the two treatment regimens (applying vs. not applying ultrasound). The alternative hypothesis was that there was a difference in the probability of achieving the primary endpoint after 24 months of treatment between the two treatment regimens (applying vs. not applying ultrasound).

Statistical decision rule

The ARCTIC protocol was designed to address a single primary endpoint. Statistical significance was claimed if the null hypothesis was rejected at the significance level (alpha) of 0.05 (two-sided).

3.4.10 Analyses sets in the ARCTIC trial

In the following section the analyses sets in the ARCTIC trial are defined.

Enrolled

The Enrolled set included all patients who have provided informed consent and had been included into the study database.

Full Analysis Set

The Full Analysis Set (FAS) was defined as all patients randomly assigned to a treatment group, and had started the allocated intervention defined as having completed at least one regular visit after the baseline visit. The FAS formed the primary analysis set of the study, and was used for all primary and secondary endpoints.

Safety Analysis Set

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

Completer Analysis Set

The Completer Analysis Set included all randomised patients having started the allocated intervention and not withdrawn during the study.

Per Protocol Analysis Set

The Per Protocol Analysis Set (PPS) included all randomised patients meeting the study entry criteria and with no major protocol deviations.

3.4.11 Missing data

In paper I, a completer analysis approach was used to assess response, meaning that patients with missing data at the selected point of assessment were not included in the analysis. This approach is likely to overestimate effectiveness, as patients with poorer response are more likely to discontinue treatment early. In paper III and IV we also used a completer analysis approach.

In paper II, we used a conservative approach for missing data. For the primary endpoint (consisting of radiographic scores, DAS and 44 swollen joints) missing data were handled as follows:

Radiographic scores

- If the radiographic score were missing at month 24, the patient was considered not in complete DAS remission (failure)
- If a radiographic score was missing for visit 11 (16 months), we used last radiographic observation

DAS

- If unable to calculate DAS at visit 13 (month 24), the patient was considered not in complete DAS remission (failure)
- If unable to calculate DAS at visit 11 or 12 (month 16/20), we used last DAS observation

SJC44

- If SJC44 at visit 13 (month 24) was missing, the patient was considered not in complete DAS remission (failure)
- If SJC44 was missing at visit 11 or 12 (month 16/20), we used last SJC44 observation

3.5 Legal and ethical aspects

All studies were conducted according the principles of the Declaration of Helsinki. (155). All patients provided written, informed consent before inclusion. Each study was evaluated and approved by the Regional Ethics Committee of Eastern Norway.

3.6 Good clinical practice and monitoring

The ARCTIC trial followed the principles outlined in “Guidelines for Good Clinical Practice” (GCP) developed by International Conference on Harmonisation, Tripartite Guideline (January 1997). GCP is an international ethical and scientific quality standard for designing, conducting, recording and reporting trials that involve the participation of human subjects. The ARCTIC trial was monitored by Section for GCP, Oslo University Hospital and Innovest AS. Monitoring is the act of overseeing the progress of a clinical trial, and ensuring that it is conducted, recorded, and reported in accordance with the protocol, Standard Operating Procedures (SOPs), GCP, and the applicable regulatory requirements.

4. Summary of results

4.1 Paper I

Time trends in disease activity, response and remission rates in rheumatoid arthritis during the past decade: results from the NOR-DMARD study 2000–2010

The main objective of this study was to investigate whether baseline disease activity levels and treatment responses in patients with RA changed during the period 2000–2010. Data were provided by the NOR-DMARD study. A total of 1866 patients starting MTX monotherapy and 707 patients starting their first TNFi+MTX regimen were included consecutively and followed longitudinally.

Significant time trends towards lower baseline DAS28 as well as other disease activity measures were found in both groups (mean DAS28 was reduced from 5.17 to 4.75 in the MTX monotherapy group, and from 5.88 to 4.64 in TNFi+MTX group), and disease duration became shorter. Six-month DAS28 remission rates increased significantly over the years (from 17.8% to 37.6% for MTX monotherapy, from 16.9% to 46.3% for TNFi+MTX). MTX doses increased significantly in both treatment groups during the decade. The proportion of patients using co-medication with prednisolone at initiation of MTX monotherapy increased over the years, while the proportion still using prednisolone after 6 months decreased. Two-year drug survival for MTX monotherapy improved over the decade, while there was no significant trend for change in drug survival at 2 years for the TNFi+MTX group.

Baseline RA disease activity level at the time of starting MTX as well as TNFi+MTX decreased from high to moderate during the last decade. A more than two-fold increase in 6-month remission rates was observed in both groups.

4.2 Paper II

Ultrasound in the management of rheumatoid arthritis: results from the randomised controlled ARCTIC trial

Ultrasound at point of care has become widely implemented in rheumatology clinical practice, but the added value of ultrasound examinations in early RA has not been studied. The aim of the ARCTIC trial was to examine if a treatment strategy based on structured ultrasound assessment would lead to improved outcomes in RA, compared to a conventional strategy.

Patients were recruited between September 2010 and April 2013 and randomly assigned to receive ultrasound (n=122) or conventional tight control (n=116). Twenty-six of the 118 analysed patients (22.0%) in the ultrasound tight control arm and 21 of the 112 analysed patients (18.8%) in the clinical tight control arm reached the primary endpoint (mean difference 3.3%; 95%CI -7.1% to 13.7%). No significant difference was found between the groups for any of the three components of the primary outcome – no swollen joints at 16, 20 and 24 months, DAS remission at 16, 20 and 24 months, and no radiographic progression between 16 and 24 months.

At 24 months 67.8% and 67.0% of the patients in the two arms were in DAS remission. No significant differences in remission rates were found at 12 or 24 months for either of the remission criteria sets analysed. More than 80% of patients in both arms showed EULAR good or moderate response at 12 months and this proportion was sustained throughout the study.

More patients in the ultrasound tight control arm (28.8%) than in the conventional arm (17.0%) received biologic treatment at the end of the study, and fewer patients remained on MTX monotherapy (53.4% vs. 71.4%).

The systematic use of ultrasound in the follow-up of early RA patients treated according to current recommendations is not justified based on the ARCTIC results.

4.3 Paper III

First step in the development of an ultrasound joint inflammation score for rheumatoid arthritis using a data-driven approach

The main objective of this study was, by a data-driven approach in an early RA cohort, to develop feasible candidate ultrasound joint inflammation scores with most information retained (ARCTIC) and subsequently validate these scores in a cohort of patients with established active RA (ULRABIT).

227 early and 212 established RA patients were included. We identified two candidate sets of joints/tendons: Candidate set A consisted of 7 joints/2 tendons (MCP1, MCP2, PIP3, radiocarpal, elbow, MTP1, MTP2, tibialis posterior tendon, extensor carpi ulnaris tendon), and set B of 9 joints/2 tendons (set A with MCP5 and MTP5 added). Unilateral reduced scores retained 78%–85% of the information in total score, while bilateral reduced scores retained 89%–93%, and both sets performed better than previously proposed reduced joint scores in a separate validation cohort of established RA.

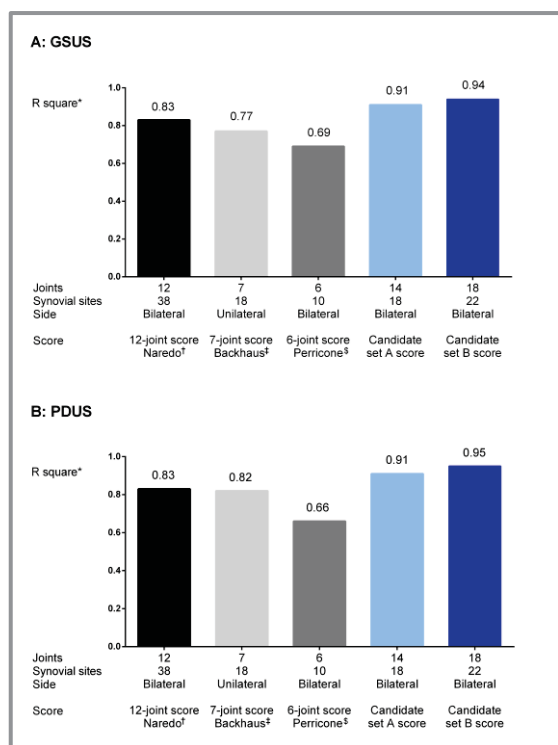


Figure: Proportion of information in the total score retained (*R square*) by previously proposed ultrasound joint scores and candidate set A and B in the established RA cohort

*Linear regression analysis with the total US score as dependent variable and the simple sum US score of the reduced joint scores as independent variable.

12-joint score Naredo et al. (108): elbow, wrist, MCP2, MCP3, knee, ankle and tibialis posterior 5endon bilaterally.

7-joint score Backhaus et al. (105): wrist, MCP2, MCP3, PIP2, PIP3, MTP2, MTP5 and extensor carpi ulnaris tendon dominant side.

6-joint score Perricone et al. (109): wrist, MCP2, knee bilaterally.

The reduced GSUS and PDUS scores retained most of the information from the total score, and performed well in a validation cohort of established RA.

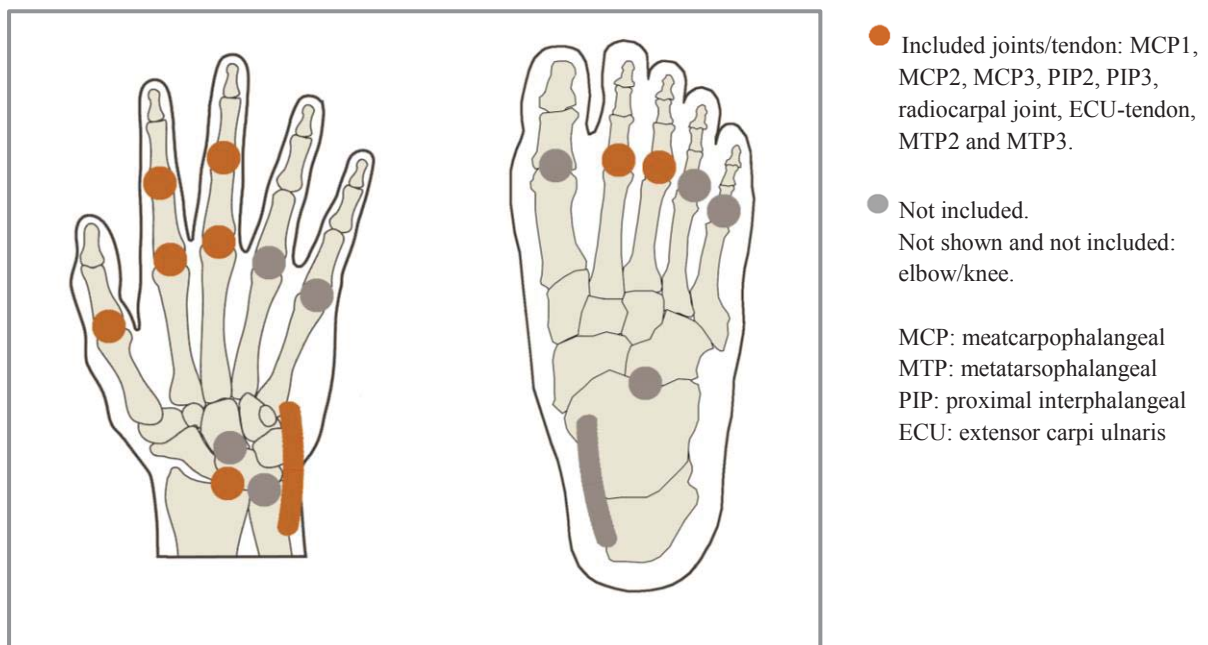
4.4 Paper IV

Development of a feasible and responsive ultrasound joint inflammation score for rheumatoid arthritis through a data-driven approach

The objective of this study was to further optimize the ultrasound joint inflammation score for RA in terms of sensitivity to change and feasibility, to validate the reduced score in an independent longitudinal cohort, and to perform comparative analyses with formerly proposed ultrasound scores.

118 early (ARCTIC) and 212 established (ULRABIT) RA patients were included. The final ultrasound score included eight joints (MCP1, MCP2, MCP3, PIP2, PIP3, radiocarpal, MTP2 and MTP3) and one tendon (ECU) examined bilaterally. The 6-month SRMs for the final score were -1.24 (95% CI -1.47 to -1.02) for GSUS and -1.09 (-1.25 to -0.92) for PDUS in early RA, with 90% of total information retained for GSUS and 92% for PDUS.

Figure 3: *The final ultrasound inflammation score for rheumatoid arthritis (USRA8)*



Our final ultrasound inflammation score (USRA8) showed good responsiveness, retained most of the information from the full score, and overall performed better than previous scores in a validation cohort.

5. Discussion

5.1 Discussion of methodological aspects

In this chapter, methodological limitations, strengths and possible sources of bias in the present thesis will be discussed.

5.1.1 Study design

The three cohorts used in this thesis represent different study designs, with different strengths and weaknesses. The broad inclusion criteria and few exclusion criteria of the observational NOR-DMARD study enables to capture the whole range of patients treated in clinical practice, including patients with lower levels of disease activity and patients with comorbidities. Additional strengths include the very long inclusion period and recruiting of patients from five different rheumatology departments in different parts of Norway. Weaknesses of this study include the lack of radiographic data, biobank material and loss to follow up.

The longitudinal observational ULRABIT study included patients starting or switching biologic DMARDs. The number of previously used biologic DMARD for the included patients ranged from zero to six, which is reflecting clinical practice, but also makes it a heterogeneous established RA cohort. The ULRABIT study included an extensive ultrasound examination at several time points, but was a single centre study, which might reduce the generalisability.

A randomised controlled trial can provide the highest category of evidence for therapeutic interventions in medicine (145). Compared to observational studies, inclusion and exclusion criteria of randomised controlled trials are stricter, they are more expensive to conduct and there is a higher risk associated with initiating such a study. We succeeded in the inclusion and follow-up of patients in the multi-centre ARCTIC trial and completed the study within a reasonable time frame, and we believe the results are generalisable to the majority of early RA patients in clinical practice.

5.1.2 Representativeness of study populations

Whether the patients and treatment practices in the NOR-DMARD centres are representative for Norwegian rheumatology as a whole is not known, but is presumable due to the broad inclusion criteria. It has been estimated that approximately 85% of patients from the participating centres who started DMARD in the study period 2000–2010 were included. The diagnosis of RA was based on clinical judgement of the treating rheumatologist.

The liberal use of biologic DMARDs in Norway may reduce the generalisability of our findings in paper I to other countries with stricter criteria for initiating biologic DMARDs. However, the two-fold increase in remission rates for MTX monotherapy illustrates that good results also can be achieved with cheap drugs if they are used early according to updated treatment recommendations. The strength of NOR-DMARD is related to its real-life setting.

The ARCTIC trial can be viewed as an inception cohort of RA patients. This study is one of the first strategy trials including patients according to the 2010 ACR/EULAR classification criteria for RA, which is known to facilitate earlier classification of RA (156). We do not know if the patients included are representative. However, there were no requirements regarding disease activity level, but a clinical indication for initiation of MTX was required, and the patients had to be DMARD naïve. The baseline findings are in line with other inception cohorts of RA (157, 158), except for a higher rate of anti-CCP positive patients, which probably is due to the new classification criteria for RA where positivity for anti-CCP and RF is strongly weighted.

A weakness of the NOR-DMARD study was that radiographic data were not systematically assessed as part of the data collection before 2012. Thus, we did not have radiographic data for patients included in paper I. Radiographic data would have contributed with important information, but we had some information regarding the proportion of patients with erosive disease at baseline, which decreased during the study period in the biologic DMARD group. Furthermore, radiographic damage is associated to physical function, particularly in established disease (159) and baseline MHAQ score also improved during the decade.

The data collection in paper I included date of diagnosis, and not the time from patient reported first swollen joint. The reduction in time from first swollen joint to initiation of methotrexate during the period 2000–2010 would probably have been even more pronounced than reduction in time from diagnosis to initiation of therapy, since the clinicians were increasingly aware of the importance of initiation of DMARD therapy as soon as the diagnosis of RA was verified. How to define disease onset differs across studies (160).

In the ARCTIC trial we calculated DAS with 44 swollen joints and Ritchie articular index (RAI), as DAS remission is a more stringent outcome than DAS28 remission (161). RAI has a different approach to assess tenderness than 28-tender joints, as joints and joint groups are scored 0–3 instead of 0 or 1 (59), and certain joints are treated as a single unit (metacarpophalangeal and proximal interphalangeal joints of each hand and the metatarsophalangeal joints of each foot). Each single joint was examined, but only the unit scores of RAI were registered in the electronic Case Report Form. In retrospect, it would have been an advantage to also record 28 tender joints (each MCP and PIP joint of the hand), for direct calculation of DAS28 which would have provided opportunities for easier comparison of our results with previous studies.

In the ARCTIC trial conventional radiographs were read according to van der Heijde-modified Sharp score (90) by two trained and blinded readers at a central facility, with known chronological order, since it is found to be the most sensitive method for detecting changes over time (87, 162). An implication of this sensitive scoring method of conventional radiographs is that a higher proportion of patients in both groups of the ARCTIC trial were considered to progress radiographically, than would be revealed with blinded chronology. It is known that the van der Heijde-modified Sharp scoring system is more time consuming than the Larsen method or its modifications. However, for trials where small differences are important, the modified Sharp scores seem to be the most appropriate (162).

A pilot study demonstrated good intra- and inter-reader reliability for the ultrasound scoring system we used in paper II, III and IV (102), but no consensus has been reached on the

optimal scoring system for ultrasound in RA. The sonographers were experienced, and underwent extensive training and calibration during the study.

5.1.4 Statistical considerations

Analysis of completers

In the longitudinal observational study in paper I response and remission data are presented for completers, meaning that only patients that had recorded outcome measures at the time point for assessment were included. This approach might overestimate the response and remission rates, as patients with poorer response are more likely to discontinue treatment early. However, the results were consistent after LUNDEX correction. The LUNDEX is a tool for evaluating drug effectiveness in observational studies. It has the advantage of integrating clinical response as well as adherence to therapy in a composite value.

(LUNDEX value = [fraction of starters still in study at time T] x [fraction responding at time T]) (163).

In the randomised controlled trial in paper II a conservative approach was used for the efficacy analysis. The primary efficacy analyses were based on the Full Analysis Set, with sensitivity analysis using the Per Protocol Analysis Set. Secondary efficacy analyses were based on the Full Analysis Set alone. As there was only one predefined primary analysis, there were no adjustments for multiple testing in the secondary analyses. Missing data were treated as worst outcome for important outcomes, for details and definitions see chapter 3.4.10 and 3.4.11. The response and remission rates and other endpoints are conservatively estimated by this approach. Over all, missing data were less than 10% in the ARCTIC trial.

Negative trials

The randomised controlled trial in paper II did not reject the primary null hypothesis of the trial: “There is no difference in the probability of achieving the primary endpoint (complete DAS remission) after 24 months of treatment between the two treatment regimens (applying vs not applying ultrasound)”. Failure to show an effect does not automatically imply a lack of effect of the intervention; the true effect might also be of insufficient magnitude large to be discovered by the trial. Negative trials can be divided into two categories: 1) True negative trials where the trial can rule out clinically important effects, and 2) Inconclusive

trials where important clinical effects cannot be ruled out. In the following section it is discussed why we suggest that the ARCTIC trial is a true negative trial.

A central aspect in the assessment of a negative study is to consider the size of a clinically important potential effect. In our sample size calculations, we aimed to power the trial at 80% to detect a 20% difference between the interventions (cf. 3.4.8 Sample size estimation). This was based on the remission rates in previous studies (see table 3), in addition to discussions with rheumatologists regarding the effect size needed in order to introduce ultrasound in clinical practice.

The estimated treatment difference of the primary endpoint was 3.3% with a 95% confidence interval of -7.1 to 13.7. The confidence interval is completely within the $\pm 20\%$ margin, ruling out a clinically important difference between the treatments according to our estimate of an important clinical effect. Also for the individual components of the primary endpoint, the 95% confidence intervals are within the $\pm 20\%$ margin, indicating that there is no clinically important difference between the two strategies with respect to these parameters (no swollen joints, DAS remission and no radiographic progression).

5.2 Discussion of main results

In the following chapter, the main results according to the specific research questions of this thesis will be interpreted, discussed and compared to other related studies.

5.2.1 Baseline disease activity level and disease duration 2000–2010

We found that the average disease activity level at initiation of therapy decreased from high to moderate from 2000 to 2010 in patients starting MTX monotherapy as well as TNFi+MTX. We also found that the time from diagnosis to initiation of MTX monotherapy decreased from median 10 months to 10 days from 2000 to 2010, and the time from diagnosis to initiation of the first TNFi also decreased during the decade, from median 8 to 3.8 years.

In contrast to our study, Hetland et al. showed that the disease duration at initiation of TNFi remained unchanged in the DANBIO register during the study period from 2000 to 2005 (156), while in a study from British Society for Rheumatology Biologic Register (BSRBR)

Hyrich et al. observed a significant trend towards use of TNFi in patients with shorter disease duration, although mean disease duration remained as long as 11.4 years in 2008 (164).

Results from a meta-analysis of 14 randomised controlled trials support that disease duration at the time of DMARD initiation was the main negative predictor of response to DMARD treatment (165). Timing the “window of opportunity” is important (160), and it has been shown that ACPA-positive RA patients with symptoms <12 weeks have less progressive disease than patients with a longer symptom duration (166). Initiation of synthetic DMARD treatment as soon as the diagnosis of RA is made is in concordance with the EULAR and ACR treatment recommendations (31, 47). The results from paper I indicate that the clinicians during the decade had become increasingly aware that all therapies – synthetic DMARD monotherapy, synthetic DMARD combinations and biologic DMARDs – seem to work better in early disease than in established RA (40-42). Thus, Norwegian rheumatologists seemed to adopt a more aggressive treatment strategy for RA with earlier use of both synthetic DMARD and biologic DMARD and with initiation of treatment at a lower disease activity level during the period 2000 to 2010. This more aggressive treatment strategy started before the development of the new ACR/EULAR classification criteria for RA, focusing on earlier diagnosis and earlier institution of disease modifying therapy.

5.2.2 Response and remission rates 2000–2010

A two-fold increase in remission rates as well as a consistent improvement in other endpoints was observed from 2000 to 2010. This increase in response and remission rates was present both for MTX monotherapy and combination therapy with TNFi.

Previous register studies have also indicated that lower disease activity at initiation of synthetic and biologic DMARD will result in improved remission rates. The baseline disease activity in the DANBIO register decreased and treatment responses improved in RA patients treated with their first biologic DMARD in the period from 2000 to 2005 (156). Similarly, the BSRBR register showed a time trend towards less severe disease at baseline and improved outcomes in RA patients starting their first TNFi during the period 2001–2008 (164). However, in both these observational studies the baseline disease activity

remained high, in contrast to the current study in which the baseline disease activity decreased from high to moderate during the 10-year study period.

Drug survival for the first TNFi+MTX was largely unaffected by the year of treatment initiation, as also observed in DANBIO (156). This finding, despite the observed increase in response and remission rates, is probably due to the fact that the treatment target was gradually set higher at the same time with increasing accessibility to a number of biologic DMARDs during the recent years. This increased accessibility could allow switching patients to an alternative biologic DMARD when the response to their current treatment was only partial. Thus the respective effects of increased treatment success on drug survival and a greater inclination to switch treatment, seems to have levelled each other out. This increased switching practice also seems to be reflected in our data as we found an increased frequency of discontinuations due to lack of efficacy during the period 2007–2010 compared to earlier periods (167). The access to switching to other biologic DMARDs in situation with inadequate response or adverse events is currently not restricted in Norway. However, tight control and “treat to target” strategies were not fully implemented in the study period, but the focus on earlier switching has increased during recent years if the pre-defined treatment target is not achieved within 3–6 months (46).

We found a significant increase in doses of MTX during the 10-year study period in both the MTX monotherapy and in the TNFi+MTX group. Sokka et al. showed increasing use of MTX in several countries since the 1980s (168). A similar evolution has also occurred in Norway, and was observed in NOR-DMARD from 2000 to 2009 (167). In a Dutch inception cohort of RA from 1989 to 2009 the proportion of patients using MTX increased from 5% to 62%, and the average dose increased from 6.7 ± 1.3 in 1989 to 16.1 ± 5.5 mg/week in 2008 (169). Recent recommendations have also advocated that MTX should be used in higher doses than was previously practiced (170). An increase in the dose of MTX was also found by Hetland et al. with an increase to higher MTX doses than was observed in our study (156). It is well documented that higher weekly doses (20–30 mg) are more efficacious than lower doses (5–7.5mg) (31, 171). Interestingly, we also found a significant time trend for improved drug survival of MTX monotherapy during the decade, despite increased availability of TNFi treatment, which probably reflects a real improved efficacy of MTX with higher doses and earlier initiation, as well as fewer discontinuations due to non-

serious adverse events. Use of NSAIDs as co-medication in RA patients has decreased in the same time period (172), and might have led to fewer cases of liver enzyme elevation leading to discontinuation of MTX. Clinicians and patients have become increasingly aware of the importance of MTX as an anchor drug in early RA (170, 173, 174). However, the comparative efficacy of combinations of MTX and other synthetic DMARDs versus MTX monotherapy is still under debate and investigation (159, 175, 176).

In our study a higher proportion of RA patients used prednisolone when starting MTX in the recent years, and an increasing proportion of patients tapered and discontinued prednisolone. This practice is in accordance with EULAR recommendations on the use of glucocorticoids in RA (177). More aggressive use of glucocorticoids is advantageous, especially for bridging the interval between initiation of DMARDs and onset of their therapeutic effect (35). Available evidence also supports that low dose or step down glucocorticoids given in addition to standard DMARD treatment can slow radiographic progression significantly in early RA (36, 178).

These new treatment strategies developed and incorporated in the management of RA, are also transferred to other areas of rheumatology like juvenile idiopathic arthritis (JIA). In a Dutch longitudinal observational study recently published, Otten et al. found biologic agents increasingly prescribed earlier in the disease of JIA and with lower disease activity, resulting in better short-term disease outcomes (179).

5.2.3 Response and remission rates 2010–2015

After two years of DMARD treatment according to current guidelines, more than two thirds of the patients with early RA were in DAS remission (paper II). Nearly half of the patients reached ACR/EULAR remission, more than half of the patients were in SDAI remission, and more than 80% achieved EULAR good/moderate response after two years. Among patients in the ARCTIC study starting their first TNFi, 42% reached ACR/EULAR remission, 55% SDAI remission and 69% DAS remission at the end of the study.

In comparison, in the longitudinal observational study presented in paper I, 14% of the patients reached ACR/EULAR remission and 23% SDAI remission after 6 months of MTX monotherapy in the time period 2009–2010. Correspondingly, among patients starting their

first TNFi in combination with MTX, 22% reached ACR/EULAR remission and 30% SDAI remission after 6 months of treatment.

The response and remission rates in early RA in the time period 2010–2015 were substantially better than the results the years before in the group of early RA patients from the NOR-DMARD study presented in paper I. Some of the reasons for the excellent improvement in response and remission rates from one year to the next were the systematic implementation of current treatment strategies including window of opportunity, aggressive DMARD escalation, bridging with prednisolone, tight control, aiming for deep remission (DAS remission and not DAS28 remission) and aiming for sustained remission. The DMARD escalation included triple combination therapy if MTX monotherapy failed, which were not often used by Norwegian clinicians the years prior the ARCTIC trial, but has shown good response and remission rates in previous studies (45, 176).

5.2.4 Feasibility of ultrasound remission as treatment target in early rheumatoid arthritis

In an ideal setting, RA remission should represent absence of inflammation and no progression of joint damage, but current clinical remission criteria do not include imaging assessments. A key question has been whether treatment of subclinical inflammation would improve long-term outcomes, or if this will lead to overtreatment (82, 114, 180, 181).

When the study protocol of the ARCTIC trial was planned, several clinicians and researchers doubted the possibility to achieve the strict treatment goal of ultrasound imaging remission with no power Doppler signal in any of the assessed joints. It was claimed that the threshold of no power Doppler signal was not an achievable goal in RA patients, as it had been shown in a number of studies that a high proportion of patients still displayed a significant amount of power Doppler activity, even if they reached clinical remission (137, 138, 182, 183). The results from the ARCTIC trial show that ultrasound imaging remission as a treatment target in early RA was feasible, when current treatment strategies were systematically implemented, and that it was reached by a majority of the patients in both treatment groups (Paper II).

5.2.5 Effect of ultrasound guided treatment of early rheumatoid arthritis

In the ARCTIC study, we assessed the benefit of adding ultrasound information to the treatment decisions and treatment target in early RA. The study was designed based on the increasing application of ultrasound in clinical practice, as well as several studies showing ultrasound examinations to improve assessment of synovitis and prediction of important patient outcomes (114). Adding ultrasound to tight control strategy was not superior to conventional tight control strategy alone with regards to the primary endpoint (sustained remission, absence of radiographic damage and absence of swollen joints between month 16 and 24 of the study). Clinical remission rates in both study arms in paper II were excellent and compared well to the results from other strategy trials in early RA (45, 50, 129-131, 184).

In both study arms there were minimal radiographic progression. No difference in radiographic progression between 16–24 months was found between the groups. There was a trend towards a difference in van der Heijde-modified Sharp score over 24 months of 0.45 units (95% CI -0.90 – 0, p-value 0.05) favouring the ultrasound tight control group, and these findings were confirmed in the sensitivity analyses in the completer data set. We do not know if this trend towards a difference in progression of joint damage might have displayed benefits of applying ultrasound if radiographic outcome over several years had been studied. The difference may be due to more frequent initiation of biologic drugs in the ultrasound tight control arm, which is known to inhibit radiographic progression independent of disease activity (14, 54).

We are aware of another trial (NCT00920478) designed to assess the value of ultrasound in RA. The clinical target of that study was low diseases activity and not remission, which is the preferred target in current treatment recommendations (47, 54). Preliminary data from this trial indicate similar clinical outcomes in both study arms (185). We also know about two other ongoing studies assessing the value of imaging in RA, a MRI-guided treatment strategy study in Denmark (NCT01656278) and PDUS-based strategy study in Europe (NCT02056184).

The outcome of our study is to a certain degree surprising, in that a growing body of evidence suggested that subclinical inflammation, especially power Doppler activity, was

associated with radiographic progression and flares (82, 114, 137, 138, 182, 183). Despite the somewhat more aggressive treatment in the ultrasound tight control group (more frequent use of biologic DMARDs, a higher number of intra-articular glucocorticoid injections), inflammation assessed by ultrasound was suppressed to a minimum in both study arms, with a majority of patients having no power Doppler activity in any joint after two years. A possible explanation may be that isolated subclinical inflammation in the absence of clinically detectable disease activity has minimal clinical importance, making direct visualization of power Doppler activity unnecessary.

The ARCTIC study highlights the importance of conducting randomised control trials to evaluate not only drugs, but also new technologies and treatment strategies. The implementation and systematic use of ultrasound in the follow-up of early RA patients treated with an aggressive tight control strategy is not justified based on the results of the ARCTIC trial.

5.2.6 Information retained in a reduced ultrasound inflammation score

Ultrasound examination can give valuable information about inflammation in clinical practice and research, but might be time consuming if several joints are evaluated. In paper III we wanted to identify joints that give most information about the total inflammation load in active RA. Feasibility would be increased if a reduced number of joints and tendons examined could give the same, or nearly the same, information about inflammation as a more extensive score. Our approach for selecting joint areas was different from former studies and was based on which joints contributed most relevant information.

This study included a comprehensive assessment of GSUS and PDUS in 36 joints and 4 tendons. The selection of these joints and tendons for the full score were based on previous studies by Hammer et al. with ultrasound scanning of 78 joints (113), by Naredo et al. with scanning of 44 joints (108), a study by Hammer et al. showing tendons to be frequently inflamed (147), a study by Backhaus including PIP2 and PIP3 (105), as well as experience from using ultrasound in the clinic. Ideally, our “gold standard” would have been even more comprehensive than the available 36 joints and 4 tendons, but such an ultrasound protocol was considered infeasible due to the large numbers of patients in our cohorts.

To our knowledge, paper III is the first study where an ultrasound inflammation score for RA was developed through a comprehensive data-driven approach. Using factor analysis, we identified two ultrasound joint/tendon candidate sets in a cohort of early RA which explained more of the variance in a total ultrasound score, than previously proposed ultrasound scores in a separate validation cohort of established RA.

The formerly proposed 7-joint score is unilateral while the 6- and 12- joint scores are bilateral (105, 108, 109). It is obviously less time consuming to examine a set of joints/tendons only on one side of the body. Our findings of high concordance between scores for joints/tendons on the same body side are supported by previous studies indicating that RA is not necessarily a symmetrical disease initially, and symmetry is not included as an item in the 2010 ACR/EULAR classification criteria for RA (2, 3, 186). Unilateral set A performed better than formerly proposed ultrasound joint scores, and may be sufficient in some clinical research settings. The total information retained was however higher if both sides were examined, and our analyses indicate that bilateral examination should be preferred in early disease and probably also in patients with moderate and low disease activity.

A major strength of this study was the inclusion of independent cohorts representing two clearly defined groups of patients with RA, in whom the same extensive ultrasound examination was performed using an ultrasound atlas as reference.

5.2.7 Responsiveness of the ultrasound score in paper IV and comparison to former scores

Our final ultrasound inflammation score (the Ultrasound in Rheumatoid Arthritis 8 joint score - USRA8) in paper IV had high responsiveness in an early RA cohort starting first synthetic DMARD, and performed overall better than previous scores including the candidate sets from paper III, in the validation cohort of established RA (105, 108, 109).

Some of the previously proposed scores examined joints using several projections. This is especially important when assessing erosive damage, but may also provide some additional information about the inflammatory load. However, applying a dorsal only approach of the wrist, MCP, PIP and MTP joints, as in our studies, increases feasibility, and it has been shown that this projection captures most of the inflammatory findings (187). When

assessing inflammation, applying several projections might not be worth it from a feasibility point of view, but these considerations might very well be different when it comes to evaluation of erosive damage.

We have not calculated the Ultrasound Global Synovitis Score (US-GLOSS) in paper III or IV. US-GLOSS is a composite score applied at joint level where each joint get a total value 0–3 where the highest value of GSUS or PDUS equals the individual joint score (188). We believe that GSUS and PDUS may reflect different aspects of inflammation. Several previous studies have shown that PDUS reflects active inflammation, and predicts radiographic progression and disease flare (137, 138, 180, 189). GSUS without PDUS may express previous inflammation and subsequent fibrotic change (182, 190, 191). There are indications that GSUS and PDUS are of different importance regarding prediction of joint damage, and the two ultrasound modalities may have different sensitivity to changes at different stages of RA (137, 138, 180, 192).

Standardised response means (SRMs) is the most commonly used statistics to estimate responsiveness, or sensitivity to change over time (150, 153). In paper IV we found that the SRMs did not change substantially if ultrasound examination is performed in few or many joints, e.g. unilateral vs. bilateral examination. This is in line with results from the APPRAISE study which showed PDUS scores were responsive under abatacept treatment, regardless of the number of joints examined by ultrasound (112). In general, all the composite indices (DAS, SDAI, CDAI) will be more responsive than single measures, as found in other studies, since combining measures reduces the scatter (150, 193).

We hope that our USRA8-score might be useful particularly in the management of established active RA, where it can be challenging to distinguish between joints with active synovitis and joints with long-standing destructions and secondary pain. The USRA8-score can be useful in clinical trials to assess ultrasound joint inflammation in RA. The objective measure of joint inflammation provided by ultrasound gives unique opportunities to answer research questions related to disease activity and inflammation in RA. The USRA8-score includes many of the same joints as in a previous study based on clinical judgement and observed frequency of involvement in mixed cohorts of RA (105), and we hope that these data can contribute to a consensual ultrasound joint inflammation score for RA.

6. Conclusions

6.1 Answer to research questions

According to the specific research questions of this study presented in section 2.2, we were able to draw the following conclusions:

- During the last decade (2000–2010) rheumatologists adopted a more aggressive treatment strategy for RA with earlier use of both synthetic and biologic DMARD and with initiation of treatment at a lower disease activity level. (Paper I)
- From 2000 to 2010 6-months response rates improved significantly, and the remission rates were doubled for both MTX monotherapy and for the first TNFi in combination with MTX. These findings indicate that modern aggressive treatment strategies result in improved short-term outcomes. (Paper I)
- After two years of DMARD treatment according to current guidelines, more than two-thirds of the patients with early RA were in DAS remission. This indicates that systematic implementation of current treatment strategies substantially improved the outcome for early RA patients compared to the results achieved during previous years. (Paper II)
- Implementation of ultrasound imaging remission as a treatment target in early RA was feasible and was reached by a majority of the patients. (Paper II)
- An ultrasound tight control strategy was not superior to a conventional tight control strategy with regard to achieving sustained clinical remission. Based on these results, the systematic use of ultrasound in the follow-up of early RA patients treated according to current recommendations is not justified. (Paper II)
- The median change in radiographic joint damage over 24 months was small, with no statistically significant differences between the two strategies. There was a trend towards more 24-months radiographic joint damage in the non-ultrasound group. (Paper II)

- Unilateral reduced scores developed by a data-driven approach retained 78%–86% of the information in total score, while bilateral reduced scores retained 91%–95% in a validation cohort of established RA. (Paper III)
- Our final ultrasound inflammation score (USRA8) was highly responsive in an early RA cohort starting treatment with first synthetic DMARD, and had good responsiveness in a validation cohort of established RA starting or switching biologic DMARD. (Paper IV)
- Our final ultrasound inflammation score had good sensitivity to change, retained most of the information from the full score, and overall performed better than previous scores in the established RA validation cohort. (Paper III and IV)

6.2 Clinical implications

In this thesis we have examined clinically important research questions related to disease management in early RA, including development and implementation of novel treatment strategies. This thesis provides findings that are of importance to clinicians.

Before the start of this thesis, a systematic literature search of Medline, Embase and Cochrane databases including all clinical trials testing clinical, functional or structural values of a targeted treatment approach in RA was conducted by the Treat-to-Target initiative (194). Few controlled studies investigating the value of target steered schedules in RA were identified, and study designs and treatment targets were heterogeneous. These studies supported clinical benefits of structured targeted treatment of RA. None of the identified studies applied a treatment target based on imaging.

A treatment target of clinical remission is recommended by both past and current recommendations from the Treat-to-Target initiative as well as from the EULAR and ACR treatment recommendations (31, 46, 47, 195). Several observational studies have shown that inflammation is present in RA patients who are in clinical remission, and that subclinical inflammation as assessed by sensitive imaging modalities is associated with structural progression and flare of disease activity (82, 196). Imaging remission as a treatment target in RA should only be selected as a preferred target if it can be proved that subclinical

inflammation can be treated, and that the clinical and radiographic outcomes for the patients will be improved by treating to imaging remission, compared to only treating to clinical remission applying conventional strategies.

Ultrasound at point of care has become widely implemented in rheumatology clinical practice, but the value of ultrasound examinations in early RA has not been studied. We examined whether addition of ultrasound information into strategic treatment decisions and targeting of therapy towards imaging remission (abrogation of inflammation as visualized by ultrasound) would lead to improved patient outcomes, compared to a conventional treat to target strategy targeting clinical remission. Paper II did not show any benefit of adding ultrasound information to follow-up of patients with early RA.

The implementation and systematic use of ultrasound in the management of patients with RA is not justified based on the results of the randomised controlled trial in paper II, in contrast to what has previously been the impression based on indirect evidence from observational studies. A conventional treat to target strategy targeting clinical remission provides excellent disease control, and the application of ultrasound imaging remission as a treatment target may lead to overtreatment and inefficient use of health care resources. The new national Norwegian treatment recommendations for RA will be based on the ARCTIC treatment protocol, and the lessons learned from the implementation of treat to target and tight control in early RA. When systematic ultrasound examination is performed for selected patients in clinical practice, results from paper III and IV showed that a reduced ultrasound score is feasible, responsive and retains most of the information from a full score.

The results from paper II will change current clinical practice and highlights the importance of conducting randomised controlled trials, not only to evaluate efficacy of drug therapies, but also new imaging technologies or new treatment strategies.

7. References

- (1) Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet* 2010;376(9746):1094-108.
- (2) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31(3):315-24.
- (3) 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;69(9):1580-8.
- (4) Eriksson JK, Neovius M, Ernestam S, Lindblad S, Simard JF, Askling J. Incidence of rheumatoid arthritis in Sweden: a nationwide population-based assessment of incidence, its determinants, and treatment penetration. *Arthritis Care Res (Hoboken)* 2013;65(6):870-8.
- (5) Kvien TK, Glennas A, Knudsrød OG, Smedstad LM, Mowinckel P, Forre O. The prevalence and severity of rheumatoid arthritis in Oslo. Results from a county register and a population survey. *Scand J Rheumatol* 1997;26(6):412-8.
- (6) Riise T, Jacobsen BK, Gran JT. Incidence and prevalence of rheumatoid arthritis in the county of Troms, northern Norway. *J Rheumatol* 2000;27(6):1386-9.
- (7) Uhlig T, Kvien TK, Glennas A, Smedstad LM, Forre O. The incidence and severity of rheumatoid arthritis, results from a county register in Oslo, Norway. *J Rheumatol* 1998;25(6):1078-84.
- (8) Kvien TK, Uhlig T, Odegard S, Heiberg MS. Epidemiological aspects of rheumatoid arthritis: the sex ratio. *Ann N Y Acad Sci* 2006;1069:212-22.
- (9) Uhlig T, Kvien TK. Is rheumatoid arthritis disappearing? *Ann Rheum Dis* 2005;64(1):7-10.
- (10) Nishimura K, Sugiyama D, Kogata Y, Tsuji G, Nakazawa T, Kawano S, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. *Ann Intern Med* 2007;146(11):797-808.
- (11) Wolfe F, Mitchell DM, Sibley JT, Fries JF, Bloch DA, Williams CA, et al. The mortality of rheumatoid arthritis. *Arthritis Rheum* 1994;37(4):481-94.

- (12) van Nies JA, de Jong Z, van der Helm-van Mil AH, Knevel R, Le Cessie S, Huizinga TW. Improved treatment strategies reduce the increased mortality risk in early RA patients. *Rheumatology (Oxford)* 2010;49(11):2210-6.
- (13) Aletaha D, Hawker G, Neogi T. Re:Clarification of the role of ultrasonography, magnetic resonance imaging and conventional radiography in the ACR/EULAR 2010 rheumatoid arthritis classification criteria - comment to the article by Aletaha et al. *Annals of the Rheumatic Diseases* 2011.
- (14) Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet* 2009;373(9664):659-72.
- (15) McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *N Engl J Med* 2011;365(23):2205-19.
- (16) MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum* 2000;43(1):30-7.
- (17) Karlson EW, Lee IM, Cook NR, Manson JE, Buring JE, Hennekens CH. A retrospective cohort study of cigarette smoking and risk of rheumatoid arthritis in female health professionals. *Arthritis Rheum* 1999;42(5):910-7.
- (18) Symmons DP, Bankhead CR, Harrison BJ, Brennan P, Barrett EM, Scott DG, et al. Blood transfusion, smoking, and obesity as risk factors for the development of rheumatoid arthritis: results from a primary care-based incident case-control study in Norfolk, England. *Arthritis Rheum* 1997;40(11):1955-61.
- (19) Vessey MP, Villard-Mackintosh L, Yeates D. Oral contraceptives, cigarette smoking and other factors in relation to arthritis. *Contraception* 1987;35(5):457-64.
- (20) Uhlig T, Hagen KB, Kvien TK. Current tobacco smoking, formal education, and the risk of rheumatoid arthritis. *J Rheumatol* 1999;26(1):47-54.
- (21) Saevarsdottir S, Wedren S, Seddighzadeh M, Bengtsson C, Wesley A, Lindblad S, et al. Patients with early rheumatoid arthritis who smoke are less likely to respond to treatment with methotrexate and tumor necrosis factor inhibitors: observations from the Epidemiological Investigation of Rheumatoid Arthritis and the Swedish Rheumatology Register cohorts. *Arthritis Rheum* 2011;63(1):26-36.
- (22) Saevarsdottir S, Rezaei H, Geborek P, Petersson I, Ernestam S, Albertsson K, et al. Current smoking status is a strong predictor of radiographic progression in early rheumatoid arthritis: results from the SWEFOT trial. *Ann Rheum Dis* 2015;74(8):1509-14.

- (23) Klareskog L, Padyukov L, Lorentzen J, Alfredsson L. Mechanisms of disease: Genetic susceptibility and environmental triggers in the development of rheumatoid arthritis. *Nat Clin Pract Rheumatol* 2006;2(8):425-33.
- (24) Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum* 2006;54(1):38-46.
- (25) Klockars M, Koskela RS, Jarvinen E, Kolari PJ, Rossi A. Silica exposure and rheumatoid arthritis: a follow up study of granite workers 1940-81. *Br Med J (Clin Res Ed)* 1987;294(6578):997-1000.
- (26) Sverdrup B, Kallberg H, Bengtsson C, Lundberg I, Padyukov L, Alfredsson L, et al. Association between occupational exposure to mineral oil and rheumatoid arthritis: results from the Swedish EIRA case-control study. *Arthritis Res Ther* 2005;7(6):R1296-303.
- (27) Auger I, Roudier J. A function for the QKRAA amino acid motif: mediating binding of DnaJ to DnaK. Implications for the association of rheumatoid arthritis with HLA-DR4. *J Clin Invest* 1997;99(8):1818-22.
- (28) Kamphuis S, Kuis W, de Jager W, Teklenburg G, Massa M, Gordon G, et al. Tolerogenic immune responses to novel T-cell epitopes from heat-shock protein 60 in juvenile idiopathic arthritis. *Lancet* 2005;366(9479):50-6.
- (29) Wegner N, Wait R, Sroka A, Eick S, Nguyen KA, Lundberg K, et al. Peptidylarginine deiminase from *Porphyromonas gingivalis* citrullinates human fibrinogen and alpha-enolase: implications for autoimmunity in rheumatoid arthritis. *Arthritis Rheum* 2010;62(9):2662-72.
- (30) Choy E. Understanding the dynamics: pathways involved in the pathogenesis of rheumatoid arthritis. *Rheumatology (Oxford)* 2012;51 Suppl 5:v3-11.
- (31) Smolen JS, Landewe R, Breedveld F, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* doi:10.1136/annrheumdis-2013-204573 2013.
- (32) Landewe R, van der Heijde D, Klareskog L, van Vollenhoven R, Fatenejad S. Disconnect between inflammation and joint destruction after treatment with etanercept plus methotrexate: results from the trial of etanercept and methotrexate with radiographic and patient outcomes. *Arthritis Rheum* 2006;54(10):3119-25.

- (33) Smolen JS, Han C, Bala M, Maini RN, Kalden JR, van der Heijde D, et al. Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study. *Arthritis Rheum* 2005;52(4):1020-30.
- (34) Smolen JS, van der Heijde D, Machold KP, Aletaha D, Landewé R. Proposal for a new nomenclature of disease-modifying antirheumatic drugs. *Annals of the Rheumatic Diseases* 2014;73(1):3-5.
- (35) Hoes JN, Jacobs JW, Buttgereit F, Bijlsma JW. Current view of glucocorticoid co-therapy with DMARDs in rheumatoid arthritis. *Nat Rev Rheumatol* 2010;6(12):693-702.
- (36) Kirwan JR, Bijlsma JW, Boers M, Shea BJ. Effects of glucocorticoids on radiological progression in rheumatoid arthritis. *Cochrane Database Syst Rev* 2007(1):CD006356.
- (37) Cush JJ. Early rheumatoid arthritis -- is there a window of opportunity? *J Rheumatol Suppl* 2007;80:1-7.:1-7.
- (38) Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. *Arthritis Rheum* 2003;48(7):1771-4.
- (39) Furst DE. Window of opportunity. *J Rheumatol* 2004;31(9):1677-9.
- (40) Bukhari MA, Wiles NJ, Lunt M, Harrison BJ, Scott DG, Symmons DP, et al. Influence of disease-modifying therapy on radiographic outcome in inflammatory polyarthritis at five years: results from a large observational inception study. *Arthritis Rheum* 2003;48(1):46-53.
- (41) Lard LR, Visser H, Speyer I, vander Horst-Bruinsma IE, Zwinderman AH, Breedveld FC, et al. Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. *Am J Med* 2001;111(6):446-51.
- (42) Nell VP, Machold KP, Eberl G, Stamm TA, Uffmann M, Smolen JS. Benefit of very early referral and very early therapy with disease-modifying anti-rheumatic drugs in patients with early rheumatoid arthritis. *Rheumatology (Oxford)* 2004;43(7):906-14.
- (43) Allaart CF, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Breedveld FC, Dijkmans BA. Aiming at low disease activity in rheumatoid arthritis with initial combination therapy or initial monotherapy strategies: the BeSt study. *Clin Exp Rheumatol* 2006;24(6 Suppl 43):S-82.

- (44) Verstappen SM, Jacobs JW, Van D, V, Heurkens AH, Schenk Y, ter Borg EJ, et al. Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial). *Ann Rheum Dis* 2007;66(11):1443-9.
- (45) Mottonen T, Hannonen P, Leirisalo-Repo M, Nissila M, Kautiainen H, Korpela M, et al. Comparison of combination therapy with single-drug therapy in early rheumatoid arthritis: a randomised trial. FIN-RACo trial group. *Lancet* 1999;353(9164):1568-73.
- (46) Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;75(1):3-15.
- (47) Singh JA, Saag KG, Bridges SL, Jr., Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016;68(1):1-26.
- (48) Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van ZD, Kerstens PJ, Hazes JM, et al. Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2007;146(6):406-15.
- (49) Bijlsma JW, Weinblatt ME. Optimal use of methotrexate: the advantages of tight control. *Ann Rheum Dis* 2007;66(11):1409-10.
- (50) Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;364(9430):263-9.
- (51) Saunders SA, Capell HA, Stirling A, Vallance R, Kincaid W, McMahon AD, et al. Triple therapy in early active rheumatoid arthritis: a randomized, single-blind, controlled trial comparing step-up and parallel treatment strategies. *Arthritis Rheum* 2008;58(5):1310-7.
- (52) 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;69(6):964-75.
- (53) Bakker MF, Jacobs JW, Verstappen SM, Bijlsma JW. Tight control in the treatment of rheumatoid arthritis: efficacy and feasibility. *Ann Rheum Dis* 2007;66 Suppl 3:iii56-iii60.
- (54) Smolen JS, Landewe R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic

and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73(3):492-509.

- (55) Atar D, Birkeland KI, Uhlig T. 'Treat to target': moving targets from hypertension, hyperlipidaemia and diabetes to rheumatoid arthritis. *Ann Rheum Dis* 2010;69(4):629-30.
- (56) Hetland ML, Stengaard-Pedersen K, Junker P, Lottenburger T, Ellingsen T, Andersen LS, et al. Combination treatment with methotrexate, cyclosporine, and intraarticular betamethasone compared with methotrexate and intraarticular betamethasone in early active rheumatoid arthritis: an investigator-initiated, multicenter, randomized, double-blind, parallel-group, placebo-controlled study. *Arthritis Rheum* 2006;54(5):1401-9.
- (57) Sokka T, Pincus T. Quantitative joint assessment in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;23(5 Suppl 39):S58-S62.
- (58) Sokka T, Pincus T. Joint counts to assess rheumatoid arthritis for clinical research and usual clinical care: advantages and limitations. *Rheum Dis Clin North Am* 2009;35(4):713-22, v-vi.
- (59) Ritchie DM, Boyle JA, McInnes JM, Jasani MK, Dalakos TG, Grieverson P, et al. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q J Med* 1968;37(147):393-406.
- (60) Fuchs HA, Pincus T. Reduced joint counts in controlled clinical trials in rheumatoid arthritis. *Arthritis Rheum* 1994;37(4):470-5.
- (61) Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Suppl* 1994;41:86-9.
- (62) Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. *Arthritis Rheum* 1993;36(6):729-40.
- (63) 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;23(5 Suppl 39):S100-8.
- (64) van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid

arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;49(11):916-20.

- (65) Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;7(4):R796-806.
- (66) Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford)* 2003;42(2):244-57.
- (67) Dougados M. It's good to feel better but it's better to feel good. *J Rheumatol* 2005;32(1):1-2.
- (68) Hobbs KF, Cohen MD. Rheumatoid arthritis disease measurement: a new old idea. *Rheumatology (Oxford)* 2012;51 Suppl 6:vi21-7.
- (69) Morel J, Combe B. How to predict prognosis in early rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2005;19(1):137-46.
- (70) Odegard S, Landewe R, van der Heijde D, Kvien TK, Mowinckel P, Uhlig T. Association of early radiographic damage with impaired physical function in rheumatoid arthritis: a ten-year, longitudinal observational study in 238 patients. *Arthritis Rheum* 2006;54(1):68-75.
- (71) Scott DL. Prognostic factors in early rheumatoid arthritis. *Rheumatology (Oxford)* 2000;39 Suppl 1:24-9.
- (72) Forslind K, Ahlmen M, Eberhardt K, Hafstrom I, Svensson B, Group BS. Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: role of antibodies to citrullinated peptides (anti-CCP). *Ann Rheum Dis* 2004;63(9):1090-5.
- (73) Syversen SW, Gaarder PI, Goll GL, Odegard S, Haavardsholm EA, Mowinckel P, et al. High anti-cyclic citrullinated peptide levels and an algorithm of four variables predict radiographic progression in patients with rheumatoid arthritis: results from a 10-year longitudinal study. *Ann Rheum Dis* 2008;67(2):212-7.
- (74) Skapenko A, Prots I, Schulze-Koops H. Prognostic factors in rheumatoid arthritis in the era of biologic agents. *Nat Rev Rheumatol* 2009;5(9):491-6.
- (75) Lillegraven S, Kvien TK. Measuring disability and quality of life in established rheumatoid arthritis. *Best Pract Res Clin Rheumatol* 2007;21(5):827-40.

- (76) Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: dimensions and practical applications. *Health Qual Life Outcomes* 2003;1:20.
- (77) Pincus T, Summey JA, Soraci SA, Jr., Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;26(11):1346-53.
- (78) Oude Voshaar MA, Ten Klooster PM, Glas CA, Vonkeman HE, Taal E, Krishnan E, et al. Validity and measurement precision of the PROMIS physical function item bank and a content validity-driven 20-item short form in rheumatoid arthritis compared with traditional measures. *Rheumatology (Oxford)* 2015;54(12):2221-9.
- (79) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30(6):473-83.
- (80) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21(2):271-92.
- (81) Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20(10):1727-36.
- (82) Haavardsholm EA, Lie E, Lillegraven S. Should modern imaging be part of remission criteria in rheumatoid arthritis? *Best Pract Res Clin Rheumatol* 2012;26(6):767-85.
- (83) 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;30(6):1385-6.
- (84) Haavardsholm EA, Ostergaard M, Ejbjerg BJ, Kvan NP, Uhlig TA, Lilleas FG, et al. Reliability and sensitivity to change of the OMERACT rheumatoid arthritis magnetic resonance imaging score in a multireader, longitudinal setting. *Arthritis Rheum* 2005;52(12):3860-7.
- (85) 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;66(9):1216-20.
- (86) 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;38(9):2045-50.

- (87) Boini S, Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. *Ann Rheum Dis* 2001;60(9):817-27.
- (88) Sharp JT, Young DY, Bluhm GB, Brook A, Brower AC, Corbett M, et al. How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis? *Arthritis Rheum* 1985;28(12):1326-35.
- (89) Ravindran V, Rachapalli S. An overview of commonly used radiographic scoring methods in rheumatoid arthritis clinical trials. *Clin Rheumatol* 2011;30(1):1-6.
- (90) van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 1999;26(3):743-5.
- (91) Rau R, Wassenberg S, Herborn G, Stucki G, Gebler A. A new method of scoring radiographic change in rheumatoid arthritis. *J Rheumatol* 1998;25(11):2094-107.
- (92) Szkudlarek M, Wakefield RJ, Backhaus M, Terslev L. The discriminatory capacity of ultrasound in rheumatoid arthritis: active vs inactive, early vs advanced, and more. *Rheumatology (Oxford)* 2012;51 Suppl 7:vii6-9.
- (93) Mandl P, Naredo E, Wakefield RJ, Conaghan PG, d'Agostino MA. A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. *J Rheumatol* 2011;38(9):2055-62.
- (94) Dougados M, Jousse-Joulin S, Mistretta F, d'Agostino MA, Backhaus M, Bentin J, et al. Evaluation of several ultrasonography scoring systems for synovitis and comparison to clinical examination: results from a prospective multicentre study of rheumatoid arthritis. *Ann Rheum Dis* 2010;69(5):828-33.
- (95) Dohn UM, Ejbjerg BJ, Court-Payen M, Hasselquist M, Narvestad E, Szkudlarek M, et al. Are bone erosions detected by magnetic resonance imaging and ultrasonography true erosions? A comparison with computed tomography in rheumatoid arthritis metacarpophalangeal joints. *Arthritis Res Ther* 2006;8(4):R110.
- (96) Mathiessen A, Haugen IK, Slatkowsky-Christensen B, Boyesen P, Kvien TK, Hammer HB. Ultrasonographic assessment of osteophytes in 127 patients with hand osteoarthritis: exploring reliability and associations with MRI, radiographs and clinical joint findings. *Ann Rheum Dis* 2013;72(1):51-6.
- (97) Szkudlarek M, Narvestad E, Klarlund M, Court-Payen M, Thomsen HS, Ostergaard M. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography, and clinical examination. *Arthritis Rheum* 2004;50(7):2103-12.

- (98) Tan YK, Ostergaard M, Conaghan PG. Imaging tools in rheumatoid arthritis: ultrasound vs magnetic resonance imaging. *Rheumatology (Oxford)* 2012;51 Suppl 7:vii36-42.
- (99) Szkudlarek M, Terslev L, Wakefield RJ, Backhaus M, Balint PV, Bruyn GA, et al. Summary Findings of a Systematic Literature Review of the Ultrasound Assessment of Bone Erosions in Rheumatoid Arthritis. *J Rheumatol* 2016;43(1):12-21.
- (100) Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, d'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005;32(12):2485-7.
- (101) Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003;48(4):955-62.
- (102) 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;70(11):1995-8.
- (103) Tan YK, Allen JC, Lye WK, Conaghan PG, D'Agostino MA, Chew L-C, et al. Novel Ultrasound Joint Selection Methods Using a Reduced Joint Number Demonstrate Inflammatory Improvement when Compared to Existing Methods and Disease Activity Score at 28 Joints. *The Journal of Rheumatology* 2016;43(1):34-7.
- (104) Cheung PP, Lahiri M, Teng GG, Lim AYN, Lau TC, Lateef A, et al. A randomized controlled trial for improving patient self-assessment of synovitis in rheumatoid arthritis with education by ultrasonography: the RAEUS Study. *Rheumatology* 2015;54(7):1161-9.
- (105) Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. *Arthritis Rheum* 2009;61(9):1194-201.
- (106) Ellegaard K, Torp-Pedersen S, Terslev L, neskiold-Samsøe B, Henriksen M, Bliddal H. Ultrasound colour Doppler measurements in a single joint as measure of disease activity in patients with rheumatoid arthritis--assessment of concurrent validity. *Rheumatology (Oxford)* 2009;48(3):254-7.
- (107) Iagnocco A, Filippucci E, Perella C, Ceccarelli F, Cassara E, Alessandri C, et al. Clinical and ultrasonographic monitoring of response to adalimumab treatment in rheumatoid arthritis. *J Rheumatol* 2008;35(1):35-40.

- (108) Naredo E, Rodriguez M, Campos C, Rodriguez-Heredia JM, Medina JA, Giner E, et al. Validity, reproducibility, and responsiveness of a twelve-joint simplified power doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. *Arthritis Rheum* 2008;59(4):515-22.
- (109) Perricone C, Ceccarelli F, Modesti M, Vavala C, Di FM, Valesini G, et al. The 6-joint ultrasonographic assessment: a valid, sensitive-to-change and feasible method for evaluating joint inflammation in RA. *Rheumatology (Oxford)* 2012;51(5):866-73.
- (110) Scheel AK, Hermann KG, Kahler E, Pasewaldt D, Fritz J, Hamm B, et al. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. *Arthritis Rheum* 2005;52(3):733-43.
- (111) Szkudlarek M, Court-Payen, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003;48(4):955-62.
- (112) D'Agostino MA, Wakefield RJ, Berner-Hammer H, Vittecoq O, Filippou G, Balint P, et al. Value of ultrasonography as a marker of early response to abatacept in patients with rheumatoid arthritis and an inadequate response to methotrexate: results from the APPRAISE study. *Ann Rheum Dis* 2015.
- (113) Hammer HB, Kvien TK. Comparisons of 7- to 78-joint ultrasonography scores: all different joint combinations show equal response to adalimumab treatment in patients with rheumatoid arthritis. *Arthritis Res Ther* 2011;13(3):R78.
- (114) Colebatch AN, Edwards CJ, Ostergaard M, van der Heijde D, Balint PV, d'Agostino MA, et al. EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis. *Ann Rheum Dis* 2013;72(6):804-14.
- (115) Ostergaard M, Pedersen SJ, Dohn UM. Imaging in rheumatoid arthritis--status and recent advances for magnetic resonance imaging, ultrasonography, computed tomography and conventional radiography. *Best Pract Res Clin Rheumatol* 2008;22(6):1019-44.
- (116) Boyesen P, Haavardsholm EA, van der Heijde D, Ostergaard M, Hammer HB, Sesseng S, et al. Prediction of MRI erosive progression: a comparison of modern imaging modalities in early rheumatoid arthritis patients. *Ann Rheum Dis* 2011;70(1):176-9.
- (117) Haavardsholm EA, Boyesen P, Ostergaard M, Schildvold A, Kvien TK. Magnetic resonance imaging findings in 84 patients with early rheumatoid arthritis: bone marrow oedema predicts erosive progression. *Ann Rheum Dis* 2008;67(6):794-800.

- (118) Hetland ML, Ejbjerg B, Horslev-Petersen K, Jacobsen S, Vestergaard A, Jurik AG, et al. MRI bone oedema is the strongest predictor of subsequent radiographic progression in early rheumatoid arthritis. Results from a 2-year randomised controlled trial (CIMESTRA). *Ann Rheum Dis* 2009;68(3):384-90.
- (119) McQueen FM, Benton N, Perry D, Crabbe J, Robinson E, Yeoman S, et al. Bone edema scored on magnetic resonance imaging scans of the dominant carpus at presentation predicts radiographic joint damage of the hands and feet six years later in patients with rheumatoid arthritis. *Arthritis Rheum* 2003;48(7):1814-27.
- (120) Ostergaard M, Bird P, Gandjbakhch F, Eshed I, Haugen IK, Haavardsholm EA, et al. The OMERACT MRI in Arthritis Working Group - Update on Status and Future Research Priorities. *J Rheumatol* 2015;42(12):2470-2.
- (121) Ostergaard M, Edmonds J, McQueen F, Peterfy C, Lassere M, Ejbjerg B, et al. An introduction to the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. *Ann Rheum Dis* 2005;64 Suppl 1:i3-7.
- (122) Glinatsi D, Lillegraven S, Haavardsholm EA, Eshed I, Conaghan PG, Peterfy C, et al. Validation of the OMERACT Magnetic Resonance Imaging Joint Space Narrowing Score for the Wrist in a Multireader Longitudinal Trial. *J Rheumatol* 2015;42(12):2480-5.
- (123) Conaghan PG, McQueen FM, Peterfy CG, Lassere MN, Ejbjerg B, Bird P, et al. The evidence for magnetic resonance imaging as an outcome measure in proof-of-concept rheumatoid arthritis studies. *J Rheumatol* 2005;32(12):2465-9.
- (124) Conaghan PG, Durez P, Alten RE, Burmester GR, Tak PP, Klareskog L, et al. Impact of intravenous abatacept on synovitis, osteitis and structural damage in patients with rheumatoid arthritis and an inadequate response to methotrexate: the ASSET randomised controlled trial. *Ann Rheum Dis* 2013;72(8):1287-94.
- (125) Conaghan PG, Emery P, Ostergaard M, Keystone EC, Genovese MC, Hsia EC, et al. Assessment by MRI of inflammation and damage in rheumatoid arthritis patients with methotrexate inadequate response receiving golimumab: results of the GO-FORWARD trial. *Ann Rheum Dis* 2011;70(11):1968-74.
- (126) Breedveld FC, Weisman MH, Kavanaugh AF, Cohen SB, Pavelka K, van Vollenhoven R, et al. The PREMIER study: A multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;54(1):26-37.

- (127) Genovese MC, Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, et al. Etanercept versus methotrexate in patients with early rheumatoid arthritis: two-year radiographic and clinical outcomes. *Arthritis Rheum* 2002;46(6):1443-50.
- (128) St Clair EW, van der Heijde DM, Smolen JS, Maini RN, Bathon JM, Emery P, et al. Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. *Arthritis Rheum* 2004;50(11):3432-43.
- (129) Horslev-Petersen K, Hetland ML, Junker P, Podenphant J, Ellingsen T, Ahlquist P, et al. Adalimumab added to a treat-to-target strategy with methotrexate and intra-articular triamcinolone in early rheumatoid arthritis increased remission rates, function and quality of life. The OPERA Study: an investigator-initiated, randomised, double-blind, parallel-group, placebo-controlled trial. *Ann Rheum Dis* 2014;73(4):654-61.
- (130) Hetland ML, Stengaard-Pedersen K, Junker P, Lottenburger T, Hansen I, Andersen LS, et al. Aggressive combination therapy with intra-articular glucocorticoid injections and conventional disease-modifying anti-rheumatic drugs in early rheumatoid arthritis: second-year clinical and radiographic results from the CIMESTRA study. *Ann Rheum Dis* 2008;67(6):815-22.
- (131) O'Dell JR, Mikuls TR, Taylor TH, Ahluwalia V, Brophy M, Warren SR, et al. Therapies for active rheumatoid arthritis after methotrexate failure. *N Engl J Med* 2013;369(4):307-18.
- (132) Weinblatt ME, Keystone EC, Furst DE, Moreland LW, Weisman MH, Birbara CA, et al. Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum* 2003;48(1):35-45.
- (133) Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2015.
- (134) Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;70(3):404-13.
- (135) 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;66(1):34-45.

- (136) Stoffer MA, Schoels MM, Smolen JS, Aletaha D, Breedveld FC, Burmester G, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search update. *Ann Rheum Dis* 2016;75(1):16-22.
- (137) 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;58(10):2958-67.
- (138) 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;48(9):1092-7.
- (139) McQueen FM, Ostendorf B. What is MRI bone oedema in rheumatoid arthritis and why does it matter? *Arthritis Res Ther* 2006;8(6):222.
- (140) 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;70(12):2159-62.
- (141) Gandjbakhch F, Haavardsholm EA, Conaghan PG, Ejbjerg B, Foltz V, Brown AK, et al. Determining a magnetic resonance imaging inflammatory activity acceptable state without subsequent radiographic progression in rheumatoid arthritis: results from a followup MRI study of 254 patients in clinical remission or low disease activity. *J Rheumatol* 2014;41(2):398-406.
- (142) Ostergaard M, Moller-Bisgaard S. Rheumatoid arthritis: Is imaging needed to define remission in rheumatoid arthritis? *Nat Rev Rheumatol* 2014;10(6):326-8.
- (143) A guideline developers' handbook. Edinburgh: Scottish Intercollegiate Guidelines Network, . 2011.
- (144) Institute of Medicine (IOM). *Clinical Practice Guidelines We Can Trust* 2011.
- (145) Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. *BMJ* 1999;318(7183):593-6.
- (146) Bombardier C, Maetzel A. Pharmacoeconomic evaluation of new treatments: efficacy versus effectiveness studies? *Ann Rheum Dis* 1999;58 Suppl 1:I82-5.
- (147) Hammer HB, Kvien TK. sUltrasonography shows significant improvement in wrist and ankle tenosynovitis in rheumatoid arthritis patients treated with adalimumab. *Scand J Rheumatol* 2011;40(3):178-82.

- (148) Lillegraven S, Boyesen P, Hammer HB, Ostergaard M, Uhlig T, Sesseng S, et al. Tenosynovitis of the extensor carpi ulnaris tendon predicts erosive progression in early rheumatoid arthritis. *Ann Rheum Dis* 2011;70(11):2049-50.
- (149) Torp-Pedersen ST, Terslev L. Settings and artefacts relevant in colour/power Doppler ultrasound in rheumatology. *Ann Rheum Dis* 2008;67(2):143-9.
- (150) Haavardsholm EA, Ostergaard M, Hammer HB, Boyesen P, Boonen A, van der Heijde D, et al. Monitoring anti-TNFalpha treatment in rheumatoid arthritis: responsiveness of magnetic resonance imaging and ultrasonography of the dominant wrist joint compared with conventional measures of disease activity and structural damage. *Ann Rheum Dis* 2009;68(10):1572-9.
- (151) Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *Int J Integr Care* 2002;2:e15.
- (152) Wright JG, Young NL. A comparison of different indices of responsiveness. *J Clin Epidemiol* 1997;50(3):239-46.
- (153) Cohen J. *Statistical Power Analysis for the Behavioral Sciences* (2nd Edition). New Jersey: Lawrence Erlbaum Associates, Inc.; 1988.
- (154) Landewe R, van der Heijde D. Radiographic progression depicted by probability plots: presenting data with optimal use of individual values. *Arthritis Rheum* 2004;50(3):699-706.
- (155) Rickham PP. Human Experimentation. Code of Ethics of the World Medical Association. Declaration of Helsinki. *Br Med J* 1964;2(5402):177.
- (156) Hetland ML, Lindegaard HM, Hansen A, Podenphant J, Unkerskov J, Ringsdal VS, et al. Do changes in prescription practice in patients with rheumatoid arthritis treated with biological agents affect treatment response and adherence to therapy? Results from the nationwide Danish DANBIO Registry. *Ann Rheum Dis* 2008;67(7):1023-6.
- (157) Schipper LG, Vermeer M, Kuper HH, Hoekstra MO, Haagsma CJ, Den Broeder AA, et al. A tight control treatment strategy aiming for remission in early rheumatoid arthritis is more effective than usual care treatment in daily clinical practice: a study of two cohorts in the Dutch Rheumatoid Arthritis Monitoring registry. *Ann Rheum Dis* 2012;71(6):845-50.
- (158) Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van ZD, Kerstens PJ, Hazes JM, et al. Clinical and radiographic outcomes of four different treatment

strategies in patients with early rheumatoid arthritis (the BeSt study): A randomized, controlled trial. *Arthritis Rheum* 2008;58(2 Suppl):S126-S35.

- (159) Leirisalo-Repo M, Kautiainen H, Laasonen L, Korpela M, Kauppi MJ, Kaipainen-Seppanen O, et al. Infliximab for 6 months added on combination therapy in early rheumatoid arthritis: 2-year results from an investigator-initiated, randomised, double-blind, placebo-controlled study (the NEO-RACo Study). *Ann Rheum Dis* 2012.
- (160) Raza K, Saber TP, Kvien TK, Tak PP, Gerlag DM. Timing the therapeutic window of opportunity in early rheumatoid arthritis: proposal for definitions of disease duration in clinical trials. *Ann Rheum Dis* 2012.
- (161) Landewe R, van der Heijde D, van der Linden S, Boers M. Twenty-eight-joint counts invalidate the DAS28 remission definition owing to the omission of the lower extremity joints: a comparison with the original DAS remission. *Ann Rheum Dis* 2006;65(5):637-41.
- (162) van der Heijde DM. Plain X-rays in rheumatoid arthritis: overview of scoring methods, their reliability and applicability. *Baillieres Clin Rheumatol* 1996;10(3):435-53.
- (163) Kristensen LE, Saxne T, Geborek P. The LUNDEX, a new index of drug efficacy in clinical practice: results of a five-year observational study of treatment with infliximab and etanercept among rheumatoid arthritis patients in southern Sweden. *Arthritis Rheum* 2006;54(2):600-6.
- (164) Hyrich KL, Watson KD, Lunt M, Symmons DP. Changes in disease characteristics and response rates among patients in the United Kingdom starting anti-tumour necrosis factor therapy for rheumatoid arthritis between 2001 and 2008. *Rheumatology (Oxford)* 2011;50(1):117-23.
- (165) Anderson JJ, Wells G, Verhoeven AC, Felson DT. Factors predicting response to treatment in rheumatoid arthritis: the importance of disease duration. *Arthritis Rheum* 2000;43(1):22-9.
- (166) Willemze A, van der Linden MP, le CS, Trouw LA, Toes RE, Huizinga TW, et al. The window of opportunity in ACPA-positive rheumatoid arthritis is not explained by ACPA characteristics. *Ann Rheum Dis* 2011;70(9):1697-8.
- (167) Lie E. Effectiveness of synthetic and biological disease modifying antirheumatic drugs in patients with inflammatory joint diseases: University of Oslo; 2012.
- (168) Sokka T. Increases in use of methotrexate since the 1980s. *Clin Exp Rheumatol* 2010;28(5 Suppl 61):S13-S20.

- (169) Kievit W, Fransen J, de Waal Malefijt MC, den Broeder AA, van Riel PL. Treatment changes and improved outcomes in RA: an overview of a large inception cohort from 1989 to 2009. *Rheumatology (Oxford)* 2013;52(8):1500-8.
- (170) Visser K, Katchamart W, Loza E, Martinez-Lopez JA, Salliot C, Trudeau J, et al. Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. *Ann Rheum Dis* 2009;68(7):1086-93.
- (171) Burmester GR, Kivitz AJ, Kupper H, Arulmani U, Florentinus S, Goss SL, et al. Efficacy and safety of ascending methotrexate dose in combination with adalimumab: the randomised CONCERTO trial. *Ann Rheum Dis* 2015;74(6):1037-44.
- (172) Katada H, Yukawa N, Urushihara H, Tanaka S, Mimori T, Kawakami K. Prescription patterns and trends in anti-rheumatic drug use based on a large-scale claims database in Japan. *Clin Rheumatol* 2015;34(5):949-56.
- (173) Pincus T, Yazici Y, Sokka T, Aletaha D, Smolen JS. Methotrexate as the "anchor drug" for the treatment of early rheumatoid arthritis. *Clin Exp Rheumatol* 2003;21(5 Suppl 31):S179-S85.
- (174) Visser K, van der HD. Optimal dosage and route of administration of methotrexate in rheumatoid arthritis: a systematic review of the literature. *Ann Rheum Dis* 2009;68(7):1094-9.
- (175) de Jong PH, Hazes JM, Barendregt PJ, Huisman M, van ZD, van der Lubbe PA, et al. Induction therapy with a combination of DMARDs is better than methotrexate monotherapy: first results of the tREACH trial. *Ann Rheum Dis* 2012.
- (176) Moreland LW, O'Dell JR, Paulus HE, Curtis JR, Bathon JM, St Clair EW, et al. A randomized comparative effectiveness study of oral triple therapy versus etanercept plus methotrexate in early aggressive rheumatoid arthritis: the treatment of early aggressive rheumatoid arthritis trial. *Arthritis Rheum* 2012;64(9):2824-35.
- (177) Gorter SL, Bijlsma JW, Cutolo M, Gomez-Reino J, Kouloumas M, Smolen JS, et al. Current evidence for the management of rheumatoid arthritis with glucocorticoids: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis. *Ann Rheum Dis* 2010;69(6):1010-4.
- (178) Bijlsma JW. Disease control with glucocorticoid therapy in rheumatoid arthritis. *Rheumatology (Oxford)* 2012;51 Suppl 4:iv9-13.

- (179) Otten MH, Anink J, Prince FHM, Twilt M, Vastert SJ, ten Cate R, et al. Trends in prescription of biological agents and outcomes of juvenile idiopathic arthritis: results of the Dutch national Arthritis and Biologics in Children Register. *Annals of the Rheumatic Diseases* 2015;74(7):1379-86.
- (180) Wakefield RJ, D'Agostino MA, Naredo E, Buch MH, Iagnocco A, Terslev L, et al. After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? *Postgrad Med J* 2012;88(1042):482-6.
- (181) Moller-Bisgaard S, Horslev-Petersen K, Ejbjerg BJ, Boesen M, Hetland ML, Christensen R, et al. Impact of a magnetic resonance imaging-guided treat-to-target strategy on disease activity and progression in patients with rheumatoid arthritis (the IMAGINE-RA trial): study protocol for a randomized controlled trial. *Trials* 2015;16:178.
- (182) 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;70(1):172-5.
- (183) 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;71(8):1316-21.
- (184) 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* 2013;72(6):844-50.
- (185) Dale J, Stirling A, McInnes I, Porter D. Targeting Ultrasound Remission In Early Rheumatoid Arthritis – Results Of The Taser Study. *Arthritis & Rheumatism* 2013;65 (Suppl).
- (186) Funovits J, Aletaha D, Bykerk V, Combe B, Dougados M, Emery P, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: methodological report phase I. *Ann Rheum Dis* 2010;69(9):1589-95.
- (187) Witt MN, Mueller F, Weinert P, Nigg AP, Reindl CS, Proft F, et al. Ultrasound of synovitis in rheumatoid arthritis: advantages of the dorsal over the palmar approach to finger joints. *J Rheumatol* 2014;41(3):422-8.
- (188) Iagnocco A, Naredo E, Wakefield R, Bruyn GA, Collado P, Jousse-Joulin S, et al. Responsiveness in Rheumatoid Arthritis. A Report from the OMERACT 11 Ultrasound Workshop. *J Rheumatol* 2014;41(2):379-82.

- (189) 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;64(1):67-76.
- (190) Koski JM, Saarakkala S, Helle M, Hakulinen U, Heikkinen JO, Hermunen H. Power Doppler ultrasonography and synovitis: correlating ultrasound imaging with histopathological findings and evaluating the performance of ultrasound equipments. *Ann Rheum Dis* 2006;65(12):1590-5.
- (191) Saleem B, Brown AK, Keen H, Nizam S, Freeston J, Wakefield R, et al. Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. *Ann Rheum Dis* 2011;70(5):792-8.
- (192) Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P, et al. Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. *Arthritis Rheum* 2007;57(1):116-24.
- (193) Verhoeven AC, Boers M, van Der Linden S. Responsiveness of the core set, response criteria, and utilities in early rheumatoid arthritis. *Ann Rheum Dis* 2000;59(12):966-74.
- (194) Schoels M, Knevel R, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas DT, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. *Ann Rheum Dis* 2010;69(4):638-43.
- (195) 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;69(4):631-7.
- (196) van der Heijde D. Remission by imaging in rheumatoid arthritis: should this be the ultimate goal? *Annals of the Rheumatic Diseases* 2012;71(Suppl 2):i89-i92.

8. Papers I-IV

Ultrasound in the management of rheumatoid arthritis: ARCTIC randomised controlled strategy trial

Espen A. Haavardsholm, consultant rheumatologist and senior researcher¹, Anna-Birgitte Aga, consultant rheumatologist¹, Inge Christoffer Olsen, statistician¹, Siri Lillegraven, trainee in rheumatology and senior researcher¹, Hilde B. Hammer, consultant rheumatologist and senior researcher¹, Till Uhlig, professor of rheumatology¹, Hallvard Fremstad, consultant rheumatologist², Tor Magne Madland, consultant rheumatologist³, Åse Stavland Lexberg, consultant rheumatologist⁴, Hilde Haukeland, consultant rheumatologist⁵, Erik Rødevand, consultant rheumatologist and head of department of rheumatology⁶, Christian Høili, consultant rheumatologist⁷, Hilde Stray, consultant rheumatologist⁸, Anne Noraas, consultant rheumatologist⁹, Inger Johanne Widding Hansen, consultant rheumatologist¹⁰, Gunnstein Bakland, consultant rheumatologist and head of department of rheumatology^{11,12}, Lena Bugge Nordberg, research fellow¹, Désirée van der Heijde, professor of rheumatology and epidemiologist^{13,1}, Tore K. Kvien professor of rheumatology and head of department of rheumatology¹

Author Affiliations:

¹Department of Rheumatology, Diakonhjemmet Hospital, Box 23 Vinderen, 0319 Oslo, Norway

²Department of Rheumatology, Ålesund Hospital, Helse Møre og Romsdal HF, Box 1600, 6026 Ålesund, Norway

³Department of Rheumatology, Haukeland University Hospital, Helse Bergen HF, Box 1400, 5021 Bergen, Norway

⁴Department of Rheumatology, Drammen Hospital, Vestre Viken HF, Box 800, 3004 Drammen, Norway

⁵Department of Rheumatology, Martina Hansens Hospital AS, Box 823, 1306 Sandvika, Norway

⁶Department of Rheumatology, St Olavs Hospital, Trondheim University Hospital, Box 3250 Sluppen, 7006 Trondheim, Norway

⁷Department of Rheumatology, Hospital Østfold HF Moss, Box 300, 1714 Grålum, Norway

⁸Haugesund Rheumatism Hospital AS, Box 2175, 5504 Haugesund, Norway

⁹The Rheumatology Clinic Dovland/Bendvold, Tollbodgata 4, 5611 Kristiansand, Norway

¹⁰Department of Rheumatology, Sørlandet Hospital HF, Box 416, 4604 Kristiansand, Norway

¹¹Department of Rheumatology, University Hospital of North Norway, 9038 Tromsø, Norway

¹²Institute of Clinical Medicine, University of Tromsø, 9037 Tromsø, Norway

¹³Department of Rheumatology, Leiden University Medical Center, Box 9600, 2300 RC Leiden, The Netherlands

Corresponding author:

Espen A. Haavardsholm, MD PhD

Department of Rheumatology, Diakonhjemmet Hospital, Box 23 Vinderen, 0319 Oslo, Norway

E-mail: e.a.haavardsholm@medisin.uio.no, Phone: +47 99798329

Keywords: Rheumatoid arthritis, ultrasound, remission, treat-to-target

Word count: 3350

Abstract

Objectives

Application of modern imaging in clinical practice is growing. The ARCTIC trial examined if a treatment strategy based on structured ultrasound assessment would lead to improved outcomes in rheumatoid arthritis (RA), compared to a conventional strategy.

Design

Multicentre, open-label, two arm, parallel group, randomised controlled strategy trial.

Setting

Ten rheumatology departments and one specialist centre in Norway, from September 2010 to September 2015.

Participants

238 patients were recruited between September 2010 and April 2013, of which 230 (141(61%) female) received the allocated intervention and were analysed for the primary outcome. The main inclusion criteria were age of 18-75 years, fulfilment of the 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for RA, disease-modifying anti-rheumatic drug (DMARD) naïveté with indication for DMARD therapy and time from first patient reported swollen joint less than two years. Patients with abnormal renal or liver function or major co-morbidities were excluded.

Interventions

122 patients were randomised to an ultrasound tight control strategy targeting clinical and imaging remission and 116 patients were randomised to a conventional tight control strategy targeting clinical remission. Randomisation was computer-generated with block design, and allocation by sequentially numbered sealed envelopes. Patients in both arms were treated according to the same DMARD escalation strategy, with 13 visits over 2 years.

Main outcome measures

The primary endpoint was the proportion of patients with a combination between 16 and 24 months of: 1) Clinical remission 2) No swollen joints and 3) Non-progression of radiographic joint damage. Secondary outcomes included measures of disease activity, radiographic progression, functioning, quality of life and adverse events. All participants who attended at least one follow-up visit were included in the full analysis set.

Results

26 of the 118 analysed patients (22.0%) in the ultrasound tight control arm and 21 of the 112 analysed patients (18.8%) in the clinical tight control arm reached the primary endpoint (mean difference 3.3%; 95%CI -7.1% to 13.7%). Secondary endpoints (disease activity, physical function and joint damage) were similar between the two groups. Six (5.1%) patients in the ultrasound tight control arm and seven (6.3%) patients in the conventional arm experienced serious adverse events.

Conclusions

The systematic use of ultrasound in the follow-up of early RA patients treated according to current recommendations is not justified based on the ARCTIC results. The findings highlight the need for randomised trials assessing the clinical application of medical technology.

Trial registration

ClinicalTrials.gov, NCT01205854.

Introduction

The rapid development of medical equipment to help physicians in their decision-making has led to fundamental changes in patient management throughout medical specialties. The implementation of new techniques in clinical practice is often based on the opportunity to assess pathologic findings thought to be of importance, but randomised clinical trials assessing the added value of new techniques on patient outcomes are often not undertaken before changing patient care.

The management of Rheumatoid Arthritis (RA) has improved greatly over the last decade with the introduction of biologic agents, tight control strategies and early disease-modifying anti-rheumatic drug (DMARD) treatment.¹⁻⁹ With the improvement in RA care, remission has become an achievable goal for a large proportion of RA patients,¹⁰⁻¹³ but studies have shown that clinical remission not necessarily exclude progression of joint damage.^{14 15} An increasing number of rheumatologists use ultrasound in the management of rheumatic diseases. Ultrasound can assess two aspects of synovitis, the morphology and quantity by grey-scale and synovial vascularity by power Doppler, and ultrasound has been a promising tool for monitoring of disease activity in RA.¹⁶⁻¹⁸ Subclinical joint inflammation visualized by ultrasound is present in a majority of RA patients in clinical remission, and is associated with radiographic progression and disease flare in these patients.¹⁹⁻²⁴ The potential importance of ultrasound in the definition of RA remission and the monitoring of RA disease activity has led to interest in the concept of imaging remission, i.e. abrogation of inflammation assessed by sensitive imaging techniques.²⁵⁻²⁹

The ARCTIC (Aiming for Remission in rheumatoid arthritis: a randomised trial examining the benefit of ultrasound in a Clinical Tight Control regimen) study was designed to compare two tight-control treatment strategies for early RA to assess whether incorporation of ultrasound information into treatment decisions as well as targeting therapy towards imaging remission would lead to

improvement in a combined outcome of sustained clinical remission, absence of swollen joints and inhibition of joint damage.

Methods

Study design

The ARCTIC trial was a 24-month randomised, open, parallel-group clinical strategy study conducted at 11 centres in Norway, four rheumatology departments at university hospitals, six regional/community hospitals and one private practice, in compliance with the Declaration of Helsinki and the International Conference on Harmonization Guidelines for Good Clinical Practice. The study protocol and consent documents were approved by an independent ethics committee (REK Sør-Øst; reference number 2010/744) and appropriate institutional review boards. The study was conducted and analysed according to the protocol and the statistical analysis plan (online supplementary).

Participants

The main inclusion criteria were age of 18-75 years, fulfilment of the 2010 American College of Rheumatology (ACR) / European League Against Rheumatism (EULAR) classification criteria for RA,³⁰ DMARD naivety, time from first patient reported swollen joint less than two years and indication for DMARD treatment. Patients with abnormal renal or liver function or major co-morbidities were excluded (full inclusion and exclusion criteria in online supplementary, section 2). All patients provided written informed consent.

Randomisation and masking

Patients were randomised 1:1 to either an “ultrasound tight control” or a “conventional tight control” strategy. The site investigators enrolled patients, and randomisation was computer-generated with block design, allocation concealment by opaque sequentially numbered sealed envelopes (prepared by a contract research organisation, Smerud Medical Research, Oslo Norway), and stratified for

study centre and anti-CCP status. Investigators and patients were aware of the allocated treatment group. The readers of the radiographs were masked for clinical information and strategy arm, while clinical assessments including joint counts were performed by unblinded study personnel.

Assessments

Patients were assessed at 0, 1, 2, 3, 4, 6, 8, 10, 12, 14, 16, 20 and 24 months. Patients in the ultrasound strategy arm were assessed by ultrasound at every visit, according to a validated scoring system of 32 joints.³¹ According to the scoring system, these 32 joints were scored semi-quantitatively as 0-3 for both grey-scale and power Doppler: metacarpophalangeal joints 1-5, radiocarpal joint, intercarpal joint, distal radioulnar joint, elbow, knee, talocrural joint and metatarsophalangeal joints 1-5 bilaterally, giving ranges from 0 to 192 for total ultrasound score, and from 0 to 96 for grey-scale and power Doppler ultrasound score. Patients in the conventional tight control arm were assessed by ultrasound yearly, but both patient and treating physician were blinded to the results.

Treatment strategies

The ultrasound tight control strategy targeted ultrasound imaging remission, defined as no ultrasound power Doppler signal in any assessed joint, in combination with clinical remission, while the clinical tight control strategy was targeted towards clinical remission only. Clinical remission was defined as Disease Activity Score (DAS) < 1.6 and no swollen joints. The treatment in both arms was escalated according to the same algorithm (online supplementary section 3). The initial therapy was methotrexate 15 mg/week increased to 20 mg/week by week five, in combination with 7 weeks of prednisolone with tapering doses from 15 mg to zero. Further steps in the treatment algorithm included methotrexate 25 mg/week, triple synthetic DMARD therapy (methotrexate, sulphasalazine, hydroxychloroquine) and biologic treatment according to guidelines (online supplementary table S1). In both arms, swollen joints were treated by intra-articular steroids, additionally any joint with power Doppler signal in the ultrasound tight control arm should be injected. All injections in the ultrasound tight control arm were guided by ultrasound.

The clinical decision to adjust medication was based on level and change in DAS. In the ultrasound arm the physician should overrule the DAS based decision and proceed to the next treatment regimen if indicated by the ultrasound score. If the patient did not respond adequately (online supplementary table S2), the physician immediately adjusted the therapy by proceeding to the next step in the treatment algorithm. Current medication was continued if the treatment response was satisfactory. Study personnel entered information on DAS values at the present and previous visit, as well as the ultrasound score at both visits if applicable, into a web page that automatically applied the treatment decisions rules in the appropriate study arm (online supplementary table S2), and the treating physician was presented with the outcome of the response assessment and whether to escalate treatment or not.

Outcomes

The predefined primary efficacy endpoint was the proportion of patients meeting the following criteria: 1) Sustained clinical remission, defined as DAS <1.6 at 16, 20 and 24 months; 2) No swollen joints at 16, 20 and 24 months; and 3) No radiographic progression (<0.5 units change in van der Heijde modified Sharp Score) between 16 and 24 months.

Radiographs of hands, wrists, and feet were obtained at 0, 3, 6, 12, 16 and 24 months. Two trained readers, blinded for clinical data and treatment strategy, scored radiographs independently in chronological order according to the van der Heijde modified Sharp score (subscores for erosions (0 to 280) and joint-space narrowing (0 to 168), total range 0 to 448, higher scores indicating more joint damage). The average of the two readings were used for all analyses.

Secondary endpoints included DAS remission, Simplified Disease Activity Index (SDAI) remission (SDAI ranges from 0 to 86, with higher scores indicating more disease activity, remission is defined as $SDAI \leq 3.3$), ACR core set outcome variables (assessment of tender and swollen joints, pain,

patient and investigator global assessment of disease activity, ESR and C-reactive protein (CRP), and fatigue VAS. We assessed physical function by the Patient-Reported Outcomes Measurement Information (PROMIS) 20-item short form (range 20 – 100, translated to a T-score with a mean of 50 and a standard deviation (SD) of 10). EuroQol-5 Dimensions (EQ-5D) was assessed with values based on UK preference weights with range from 1 (best possible health), through 0 (death) to -0.59 (worse than death). Additionally, we assessed changes from baseline in DAS and van der Heijde modified Sharp score, the proportion of patients with radiographic progression according to different cut-offs, and EULAR response rates (details in online supplementary Statistical Analysis Plan). A complete list of secondary outcomes is provided in the online supplementary appendix. Safety was evaluated by assessment of clinical and laboratory adverse events and serious adverse events were coded according to the Medical Dictionary for Regulatory Activities (MedDRA, v.17.0).

Patient involvement

No patients were involved in the development of the research question or the outcome measures, nor were they involved in the design of the study, the recruitment of study participants and the conduct of the study. The main results from the study will be disseminated to study participants in a study newsletter.

Statistical analyses

The sample size was determined based on an assumed treatment difference of 20% (45% vs 65%) in the proportion of patients reaching the primary endpoint, yielding 98 patients in each group to reach 80% power. We aimed to include 240 patients to compensate for attrition. Efficacy and safety analyses included data from all randomised patients who initiated the allocated intervention by attending at least one regular visit after randomisation (the full analyses set).

Unadjusted logistic regression was used to assess the primary and other categorical endpoints. The analyses were not adjusted for the stratification factors centre and anti-CCP status due to low cell

frequencies. If the primary endpoint components or other categorical values were missing at month 24, values were imputed with worst outcome. Missing values prior to month 24 were imputed with the last known value. Estimates of risk difference were calculated from the logistic regression parameters using the delta method to provide confidence intervals. We used median regression with 10 000 bootstrap replications to estimate the difference in median radiographic change scores; missing values were imputed using linear intra- and extrapolation. To analyse other continuous change variables, we used analysis of covariance adjusted for baseline value in addition to study centre and anti-CCP status. Estimates of treatment difference were calculated using least-square means. Missing values were handled using multiple imputations with 10 imputations drawn from the observed distribution using the Markov-chain Monte Carlo method. The statistical significance level was set at 0.05, and all significance tests were two-sided. Secondary analyses were not adjusted for multiple testing. Further statistical details are described in the Statistical Analyses Plan and online supplementary section 4. The study is registered in the ClinicalTrials.gov database (NCT01205854). A data monitoring committee was not established for this trial. We used Stata Statistical Software version 14 for all statistical analyses.

Results

Patient disposition and baseline characteristics

Between September 30, 2010, and April 30, 2013, 238 patients with early RA were included in the ARCTIC trial (figure 1). Hundred and eighteen patients received the allocated ultrasound tight control treatment strategy and 112 patients received conventional tight control strategy and were included in the primary analysis (figure 1). Of the included patients, 104 completed the ultrasound tight control arm and 100 completed the conventional tight control arm. The frequencies and reasons for not receiving allocated treatment strategy and for discontinuation during the study were similar in both arms. The two arms were overall well balanced with regard to baseline characteristics, but more

women were randomised to the ultrasound arm (71.2% vs. 50.9%, table 1). The difference in gender distribution could not be attributed to a specific study centre (online supplementary table S5).

Table 1: Baseline characteristics

	Ultrasound tight control (n=118)	Conventional tight control (n=112)
Age (years)*	50.6 (13.3)	52.3 (14.1)
Women	84 (71.2%)	57 (50.9%)
Body mass index (kg/m ²)*	25.6 (4.4)	26.1 (4.7)
Current smoker	23 (19.5%)	32 (28.6%)
Time since patient reported first swollen joint (months)*	6.8 (5.2)	7.4 (5.6)
Anti-citrullinated peptide antibody positive	93 (78.8%)	93 (83.0%)
Rheumatoid factor positive	78 (66.1%)	86 (76.8%)
DAS*	3.51 (1.19)	3.40 (1.16)
SDAI*	25.5 (12.9)	24.4 (13.3)
Patient's global assessment of disease activity (0 – 100 mm)*	51.9 (24.7)	47.5 (23.9)
Investigator's global assessment of disease activity (0 – 100 mm)*	40.9 (20.1)	40.3 (21.2)
Swollen joint count (0 – 44)*	10.9 (7.2)	10.2 (7.8)
Tender joint count (Ritchie Articular Index, 0 – 78)*	9.2 (7.9)	8.4 (6.7)
Erythrocyte sedimentation rate, mm/hr*	23.2 (18.3)	25.8 (18.9)
C-reactive protein, mg/liter*	14.5 (20.1)	16.6 (22.39)
PROMIS Physical Function (12.1 – 62.5)*	38.9 (9.2)	39.2 (8.1)
EQ-5D (-0.59 – 1.0)†	0.66 (0.16, 0.73)	0.66 (0.47, 0.73)
Fatigue VAS (0 – 100 mm)*	43.7 (28.4)	37.0 (28.6)
van der Heijde modified Sharp Score†	3.8 (1.5, 7.5)	5.0 (2.0, 10.3)
Erosion score†	2.5 (1.0, 4.0)	3.5 (1.5, 5.8)
Joint Space Narrowing score†	1.0 (0.0, 3.0)	1.0 (0.0, 4.5)
Ultrasound total score†	24 (16, 37)	27 (14, 45)
Ultrasound grey-scale score†	17 (10, 24)	19 (10, 31)
Ultrasound power Doppler†	7 (3, 13)	7 (2, 15)

Data are number of (%) patients unless stated otherwise.

DAS=Disease Activity Score, SDAI=Simplified Disease Activity Index, PROMIS=Patient-reported Outcomes Measurement Information Score Short Form v1.0 – Physical Function 20a (reported as T-scores), EQ-5D=EuroQol-5 Dimensions.

* Mean (standard deviation).

† Median (interquartile range).

Clinical outcomes

Ultrasound tight control strategy was not superior to conventional tight control strategy with regards to the primary outcome (sustained remission, absence of radiographic damage and absence of swollen joints between month 16 and 24 of the study). Twenty-six patients (22.0%) in the ultrasound tight control group and 21 patients (18.8%) in the conventional tight control group reached the primary outcome, with a difference between the groups of 3.3 % (95% confidence interval -7.1 to

13.7, p-value 0.54) and a risk ratio of 1.18 (95 % confidence interval: 0.57-1.78). No difference was found between the groups for any of the three components of the primary outcome (table 2).

At 24 months, 67.8 % and 67.0% of the patients in the two arms were in DAS remission (table 2). No significant differences in remission rates were found at 12 or 24 months for either of the remission criteria sets analysed. DAS scores, DAS remission rates, CDAI remission rates and ACR/EULAR Boolean remission rates were comparable in both study arms during the 24-month study period (figure 2A-D). More than 80% of patients in both arms showed EULAR good or moderate response at 12 months and this was sustained throughout the study. The patient and assessor evaluation of the disease activity showed substantial improvement compared to baseline, with similar results at 12 and 24 months. Patients in both groups reported similar levels of improvement in physical function. More patients in the ultrasound tight control arm (28.8%) than in the conventional tight control arm (17.0%) received biologic treatment at the end of the study, and fewer patients received methotrexate monotherapy (53.4% vs. 71.4%). In the ultrasound tight control arm, there were 167 treatment escalations and 770 intra-articular corticosteroid injections, compared to 124 treatment escalations and 548 intra-articular corticosteroid injections in the conventional tight control arm.

Radiographic outcomes

The median change in total van der Heijde modified Sharp score over 24 months was low, with no statistically significant differences between the two strategies. A borderline statistically significant difference in the 24-month change in radiographic joint damage between the groups was observed, favouring the ultrasound tight control strategy (table 2), and this is visualised in the cumulative probability plot (figure 3). The proportion of patients with radiographic progression did not differ between the two strategy arms (table 2), and similar results were found in sensitivity analyses with different cut-offs for yearly progression (≥ 0.5 units, ≥ 2.0 units, ≥ 5 units).

Table 2: Primary and secondary outcomes

	Ultrasound tight control (n=118)	Conventional tight control (n=112)	Difference (95% CI)	p- value
Primary endpoint	26 (22.0%)	21 (18.8%)	3.3 (-7.1 – 13.7)	0.54
Components of primary endpoint				
No swollen joints ‡				
At 16, 20 and 24 months	62 (52.5%)	61 (54.5%)	-1.9 (-14.8 – 11.0)	0.77
DAS remission ‡				
At 16, 20 and 24 months	64 (54.2%)	58 (51.8%)	2.5 (-10.4 – 15.4)	0.71
No radiographic progression ‡				
Between 16-24 months	49 (41.5%)	39 (34.8%)	6.7 (-5.8 – 19.2)	0.29
Measures of disease activity, physical function and quality of life				
DAS remission (DAS<1.6) †				
At 12 months	76 (64.4%)	81 (72.3%)	-7.9 (-19.9 – 4.1)	0.20
At 24 months	80 (67.8%)	75 (67.0%)	0.8 (-11.3 – 13.0)	0.89
Δ Erythrocyte sedimentation rate, mm/hr				
At 12 months*	-11.6 (16.5)	-14.9 (18.6)	0.9 (-1.2 – 3.1)	0.39
At 24 months*	-13.0 (16.8)	-13.7 (17.7)	-1.1 (-3.3 – 1.0)	0.30
Δ Investigator's global assessment				
At 12 months*	-34.8 (19.9)	-29.9 (21.2)	-1.3 (-3.3 – 0.7)	0.19
At 24 months*	-33.0 (21.6)	-28.0 (23.2)	-2.6 (-5.5 – 0.4)	0.09
Δ Patient's global assessment				
At 12 months*	-35.1 (25.3)	-29.2 (28.8)	-1.3 (-6.3 – 3.7)	0.61
At 24 months*	-35.9 (25.9)	-29.7 (28.1)	-0.75 (-6.0 – 4.5)	0.78
SDAI remission (SDAI<3.3) †				
At 12 months	67 (56.8%)	61 (54.5%)	2.3 (-10.5 – 15.2)	0.72
At 24 months	71 (60.2%)	56 (50.0%)	10.2 (-2.6 – 23.0)	0.12
EULAR good/moderate response †				
At 12 months	97 (82.2%)	93 (83.0%)	-0.8 (-10.6 – 9.0)	0.87
At 24 months	98 (83.1%)	90 (80.4%)	2.7 (-7.3 – 12.7)	0.60
No tender joints †				
At 12 months	57 (48.3%)	64 (57.1%)	-8.8 (-21.7 – 4.0)	0.18
At 24 months	67 (56.8%)	60 (53.6%)	3.2 (-9.6 – 16.1)	0.63
No swollen joints †				
At 12 months	87 (73.7%)	78 (69.6%)	4.1 (-7.6 – 15.7)	0.49
At 24 months	84 (71.2%)	75 (67.0%)	4.2 (-7.7 – 16.2)	0.49
Δ PROMIS Physical Function				
At 12 months*	11.6 (9.1)	11.6 (8.6)	-0.3 (-2.4 – 1.8)	0.80
At 24 months*	13.0 (9.5)	11.0 (9.6)	1.28 (-1.1 – 3.6)	0.28
Δ EQ-5D				
At 12 months*	0.28 (0.28)	0.25 (0.29)	0 (-0.05 – 0.05)	0.95
At 24 months*	0.30 (0.28)	0.26 (0.28)	-0.01 (-0.06 – 0.05)	0.84
Δ Fatigue VAS				
At 12 months*	-20.0 (31.4)	-15.6 (28.5)	1.0 (-4.9 – 6.8)	0.75
At 24 months*	-20.8 (32.7)	-14.8 (28.7)	1.7 (-4.8 – 8.2)	0.61
Radiographic joint damage				
Δ Modified Sharp score at 24 months†	1.0 (0, 2.5)	1.5 (0.5, 3.0)	-0.45 (-0.90 – 0)	0.05
Δ Erosion score at 24 months†	0.5 (0, 1.5)	1.0 (0.5, 2.0)	-0.38 (-0.76 – 0.01)	0.06
Δ Joint space narrowing score at 24 months†	0.0 (0.0, 0.5)	0.0 (0.0, 0.5)	0 (-0.05 – 0.05)	>0.99
Progression, modified Sharp score >= 1.0 units/year †	43 (36.4%)	53 (47.3%)	-10.9 (-23.6 – 1.8)	0.09

	Ultrasound tight control (n=118)	Conventional tight control (n=112)	Difference (95% CI)	p- value
Ultrasound outcomes				
No power Doppler signal in any joint †				
At 12 months	94 (79.7%)	60 (53.6%)	26.1 (14.3 – 37.8)	<0.001
At 24 months	89 (75.4%)	69 (61.6%)	13.8 (1.9 – 25.7)	0.02
Medication				
DMARD regimen at 24 months				
<i>MTX monotherapy</i>	63 (53.4%)	80 (71.4%)	-18.0 (-30.3 – -5.8)	0.004
<i>MTX/SSZ/HCQ combination therapy</i>	21 (17.8%)	13 (11.6%)	6.2 (-2.9 – 15.3)	0.18
<i>Biologic treatment</i>	34 (28.8%)	19 (17.0%)	11.8 (1.1 – 22.6)	0.03
Any intraarticular injections	103 (87.3%)	86 (76.8%)	10.5 (0.6 – 20.4)	0.04
Number of intraarticular injections*	6.5 (6.0)	4.9 (5.9)	1.7 (0.2 – 3.2)	0.03
Total triamcinolone hexacetonide dose, mg* ¶	94 (88)	69 (129)	25 (-4 – 54)	0.09

Data are number (%) of patients, unless stated otherwise.

DAS=Disease Activity Score, SDAI=Simplified Disease Activity Index, EULAR= European League against Rheumatism, PROMIS=Patient-reported Outcomes Measurement Information Score Short Form v1.0 – Physical Function 20a (reported as T-scores), EQ-5D=EuroQol-5 Dimensions.

MTX/SSZ/HCQ=methotrexate/salazopyrine/hydroxychloroquine.

* Mean (standard deviation).

† Median (interquartile range).

‡ Missing data before 24 months imputed using last observation carried forward, and missing data at 24 months imputed using worst outcome.

§ Missing data were imputed using worst outcome imputation.

¶ Cumulative dose per patient.

Sensitivity analyses

In sensitivity analyses, the primary endpoint, components of the primary endpoint and the 24-month change in radiographic damage were analysed adjusted for gender (table S3), with similar results as for the main analysis. We also analysed the same variables in the completer dataset consisting of 204 patients (table S4). In these analyses, a significant difference in radiographic damage over 24 months was found, with a difference in change of van der Heijde Sharp score of 0.45 units (95% CI -0.86 to -0.39, p-value 0.03) favouring the ultrasound tight control group.

Adverse events

The overall frequency of adverse events and serious adverse events were similar in the two arms (table 3). Thirteen patients (5.7%) experienced serious adverse events, of these five (2.2%) had serious infections (table 3, online supplementary table S6). Five cases of cancer were reported: Two (follicle center lymphoma, liver metastases) in the ultrasound tight control arm and three (basal cell

carcinoma, breast cancer, squamous cell carcinoma) in the conventional tight control arm (online supplementary section 5). One patient in the ultrasound tight control arm died from pneumocystis jirovecii pneumonia (online supplementary section 6).

Table 3: Safety data from months 0 to 24

	Ultrasound tight control (n=118)	Conventional tight control (n=112)
Adverse events	417	455
Patients with adverse events	96 (81.4%)	91 (81.3%)
Patients with serious adverse events	6 (5.1%)	7 (6.3%)
Patients with serious infection	3 (2.5%) *	2 (1.8%) †
Discontinuation of study due to adverse event	7 (5.9%)	5 (4.5%)
Patients with cancer	2 (1.7%) ‡	3 (2.7%) §
Death	1 (0.8%) ¶	0

Data are number (%) of patients. *Pneumocystis jirovecii pneumonia, pneumonia and bacterial abscess. †Abscess and localised infection. ‡Follicle center lymphoma and metastases to liver. §Basal cell carcinoma, breast cancer and squamous cell carcinoma. ¶The patient died from pneumocystis jirovecii pneumonia.

Discussion

In this study, we assessed the benefit of adding ultrasound information to the treatment decisions and treatment target in early RA. The study was designed based on the increasing application of ultrasound in clinical practice, as well as several studies showing ultrasound examinations to improve assessment of synovitis and prediction of important patient outcomes.²⁴ Our study found no additional effect of an ultrasound tight control strategy compared to a conventional tight control strategy for the primary outcome of the study or for other measures of disease activity, joint damage and physical function.

The study has limitations as well as strengths. This was an open study, and two components of the primary endpoint were not blinded. Although remission rates were excellent, fewer patients than expected in the power calculations reached the strict composite primary outcome. However, the

study was sufficiently powered to show a meaningful difference, as illustrated by the confidence limits of the primary efficacy outcome excluding a clinically significant effect of the intervention (online supplementary, section 7). A pilot study demonstrated good intra- and inter reader reliability for the ultrasound scoring system,³¹ but no consensus has been reached on the optimal scoring system for ultrasound in RA.¹⁸ Our findings are strengthened by both strategies adhering to international recommendations,^{10 13} and that the primary endpoint comprising sustained clinical remission and halt of radiographic progression, thus capturing the most important aspects of RA.⁴ Two trained and blinded readers at a central facility read radiographs with known chronological order, to optimize sensitivity to change.³²

The introduction of biologic therapies in combination with aggressive treatment strategies has led to levels of disease control previously thought impossible in RA, and remission has become a realistic goal of therapy.^{3 4 9-13 15} Remission rates in both study arms of the current trial were excellent compared to other strategy trials of early RA, with minimal radiographic progression.^{1 33-36} In an ideal setting, RA remission should represent absence of inflammation and no progression of joint damage, but current clinical remission criteria do not include imaging assessments. A key question has been whether treatment of subclinical inflammation would improve long-term outcomes, or if this will lead to overtreatment.^{24-26 29} We are aware of one other trial (NCT00920478) designed to assess the value of ultrasound in RA. The clinical target of that study was low diseases activity and not remission, which is the preferred target in current treatment recommendations.^{11 13} Preliminary data from this trial indicate similar clinical outcomes in both study arms.³⁷

The outcome of the study is to a certain degree surprising, in that a growing body of evidence suggested that subclinical inflammation, especially power Doppler activity, was associated with radiographic progression and flares.^{19 20 22-25} The ultrasound strategy led to additional DMARD changes, with fewer patients remaining on methotrexate monotherapy and more patients receiving

biologic treatment, as well as a higher number of intra-articular corticosteroid injections. Despite the more aggressive treatment in the ultrasound tight control group, inflammation assessed by ultrasound was suppressed to a minimum in both study arms, with a majority of patients having no power Doppler activity in any joint after two years. A possible explanation may be that isolated subclinical inflammation in the absence of clinically detectable disease activity has minimal clinical importance, making direct visualization of power Doppler activity unnecessary. We do find that there is a trend towards a difference in progression of joint damage, and we do not know if a longer follow-up period would have revealed a benefit of the ultrasound strategy. The trend may be due to more frequent initiation of biologic drugs in the ultrasound tight control group, which is known to inhibit radiographic progression independent of disease activity.⁴ New drugs are commonly undergoing health economic evaluations before they are funded through national health care systems or private medical insurances, whereas such evaluations are less strictly imposed for new imaging modalities and other medical techniques before implementation into clinical care. Although we did not aim to analyse cost-effectiveness data, the lack of gain in benefits and the increased costs, time consumption and use of biologic drugs associated with the ultrasound tight control regimen would yield negative cost-benefit ratios.

The ARCTIC study highlights the importance of conducting randomised control trials to evaluate not only drugs, but also new technologies or new treatment strategies. The implementation and systematic use of ultrasound in the follow-up of early RA patients treated with an aggressive tight control strategy is not justified based on the results of the ARCTIC trial, and the result should be reflected in future recommendations and guidelines for managing patients with RA. Future studies are needed to assess the potential benefit of ultrasound in diagnosis of RA and in assessment of established RA patients.

What is already known on this topic

With recent improvements in treatment of rheumatoid arthritis (RA), clinical remission has become an achievable goal for a large proportion of RA patients, and is a defined target in current RA recommendations. Several observational studies have shown that inflammation is present in RA patients who are in clinical remission, and that subclinical inflammation as assessed by sensitive imaging modalities is associated with structural progression and flare of disease activity. The ARCTIC trial examined if a treatment strategy based on structured ultrasound assessment targeting imaging remission would lead to improved outcomes compared to a conventional strategy targeting clinical remission.

What this study adds

Adding ultrasound information into strategic treatment decisions and targeting therapy towards imaging remission (abrogation of inflammation as visualized by ultrasound) in RA did not lead to improved patient outcomes. The application of ultrasound imaging remission as a treatment target in RA may lead to overtreatment and inefficient use of health care resources, and the ARCTIC study highlights the importance of conducting randomised control trials, not only to evaluate drugs, but also new imaging technologies or new treatment strategies.

Acknowledgments

We thank all the patients who participated in the study, and the investigators and medical staff at all participating centres for their contributions to the study. We would also like to thank Ellen Moholt, Camilla Fongen and Maria K. Jonsson for their contributions to the data collection.

Contributors

EAH and ABA designed the study, recruited and enrolled participants, collected, analysed, and interpreted data, and wrote the report. ICO designed the study, analysed, and interpreted data, and wrote the report. SL and LBN analysed, interpreted data, and wrote the report. HBH designed the study, recruited participants, interpreted data, and reviewed the manuscript. TU designed the study, recruited and enrolled participants, interpreted data, and reviewed the manuscript. HF, TMM, ÅSL, HH, ER, CH, HS, ANB, IJWH and GB recruited and enrolled patients, interpreted data, and reviewed the manuscript. DvdH and TKK designed the study, interpreted data and reviewed the manuscript. All authors have approved the final draft and vouch for the accuracy and completeness of the data and analyses. EAH is the study guarantor.

Funding

This trial was supported by the Norwegian Research Council, Norwegian South-Eastern Health Region, Norwegian Women's Public Health Association, Norwegian Rheumatism Association, investigator initiated research grants from AbbVie, UCB Pharma, Pfizer Inc., MSD Norway, Roche Norway. Siemens Healthcare and GE Healthcare provided technical support regarding standardization of ultrasound equipment. The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Competing interests: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: EAH has received research funding from Pfizer, UCB, Roche, MSD and AbbVie for the submitted work, honoraria as a speaker from Pfizer, UCB, Roche, and AbbVie, honoraria for development of educational material from Pfizer, and has sat on advisory board for Pfizer that might have an interest on the submitted work in the previous 3 years. ABA has sat on advisory boards for UCB, AbbVie, and Pfizer, and honoraria for development of educational material for UCB, HBH has received honoraria as a speaker from AbbVie, Bristol-Myers Squibb, Roche, UCB Pharma, and Pfizer, HH has sat on advisory boards for UCB and AbbVie, GB has received honoraria as a speaker from AbbVie and has sat on advisory board for Pfizer, DvdH has received consultancy honoraria from AbbVie, Amgen, Astellas, AstraZeneca, Bristol-Myers Squibb, Celgene, Daiichi, Eli Lilly, Galapagos, Merck, Novartis, Pfizer, Roche, Sanofi Aventis, Janssen, and UCB, and is owner of Imaging Rheumatology, TKK has received consultancy honoraria from AbbVie, Bristol-Myers Squibb, Celltrion, Epirus, Hospira, Merck-Serono, MSD, Orion Pharma, Pfizer, and UCB that might have an interest on the submitted work in the previous 3 years. The other authors declare no conflicts of interests. None of the authors report that their spouses, partners, or children have financial relationships that may be relevant to the submitted work, or that they themselves have non-financial interests that may be relevant to the submitted work.

Ethics approval: The study was approved by an independent ethics committee (REC South-East; reference number 2010/744) and appropriate institutional review boards.

Transparency: The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Figures

Figure 1: Trial profile.

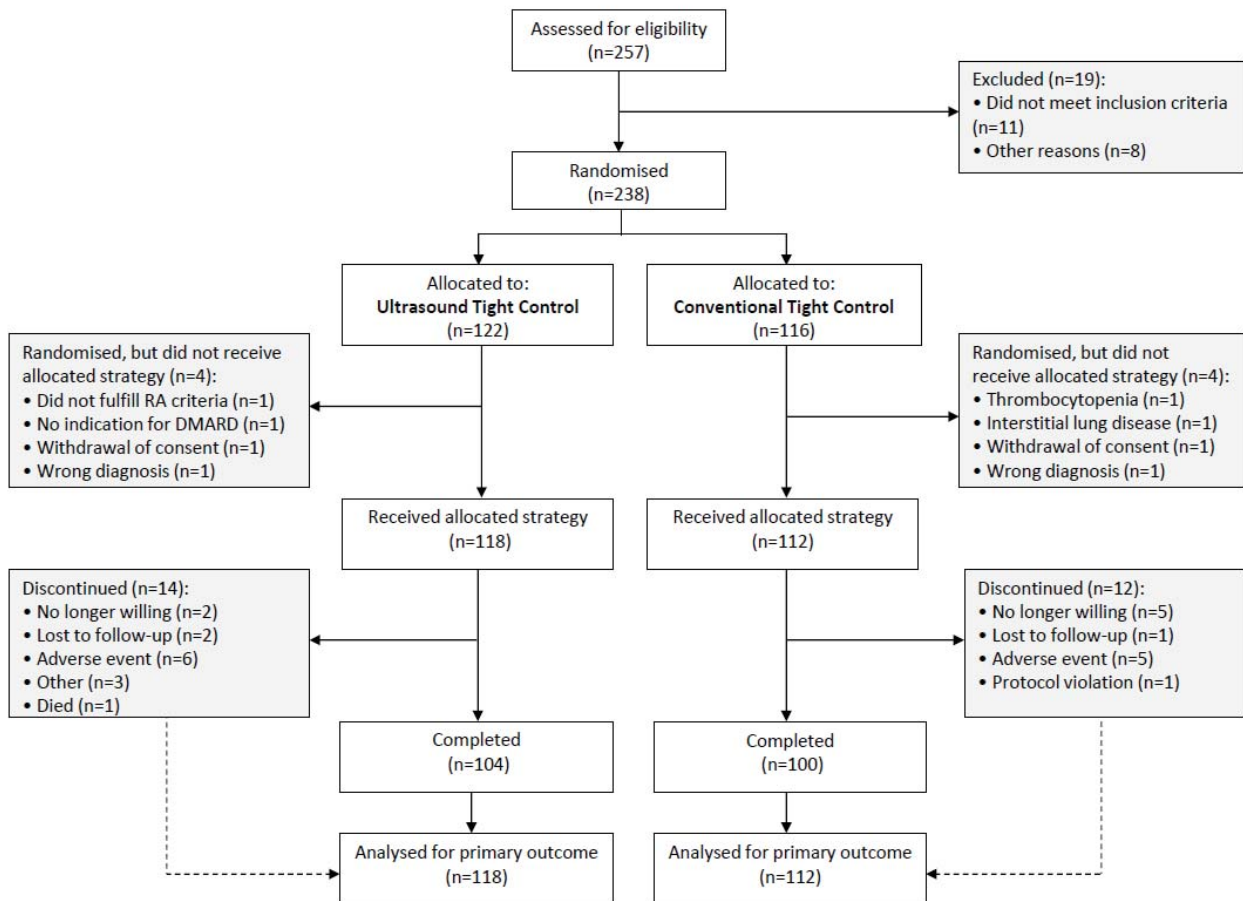


Figure 2A: Disease activity score over 24 months. Least square mean estimates of disease activity score (DAS) at all visits derived from a mixed-effects longitudinal model adjusted for baseline value and stratification factors. Bars indicate 95% confidence limits.

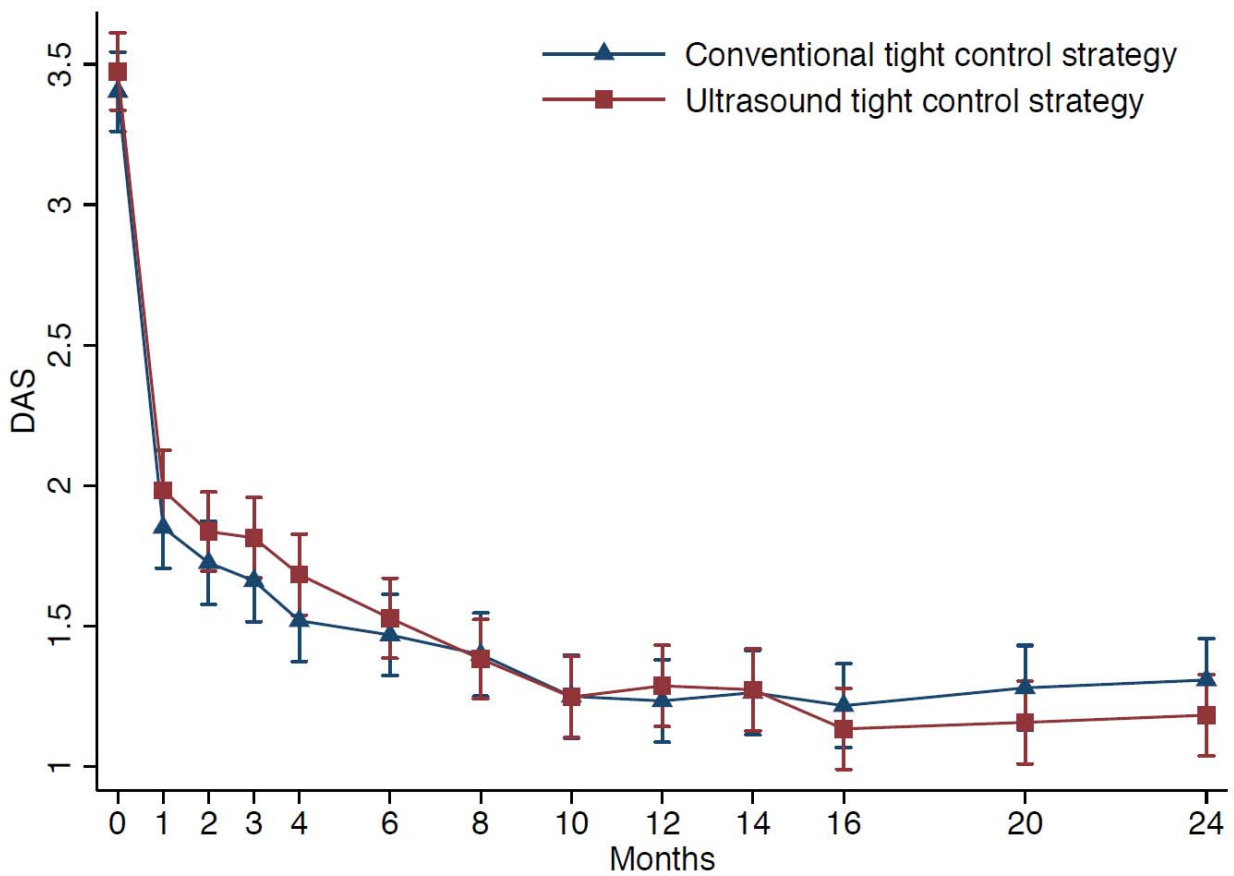


Figure 2B: Proportion of patients who achieved DAS remission over 24 months derived from a logistic regression model. Bars indicate 95% confidence limits.

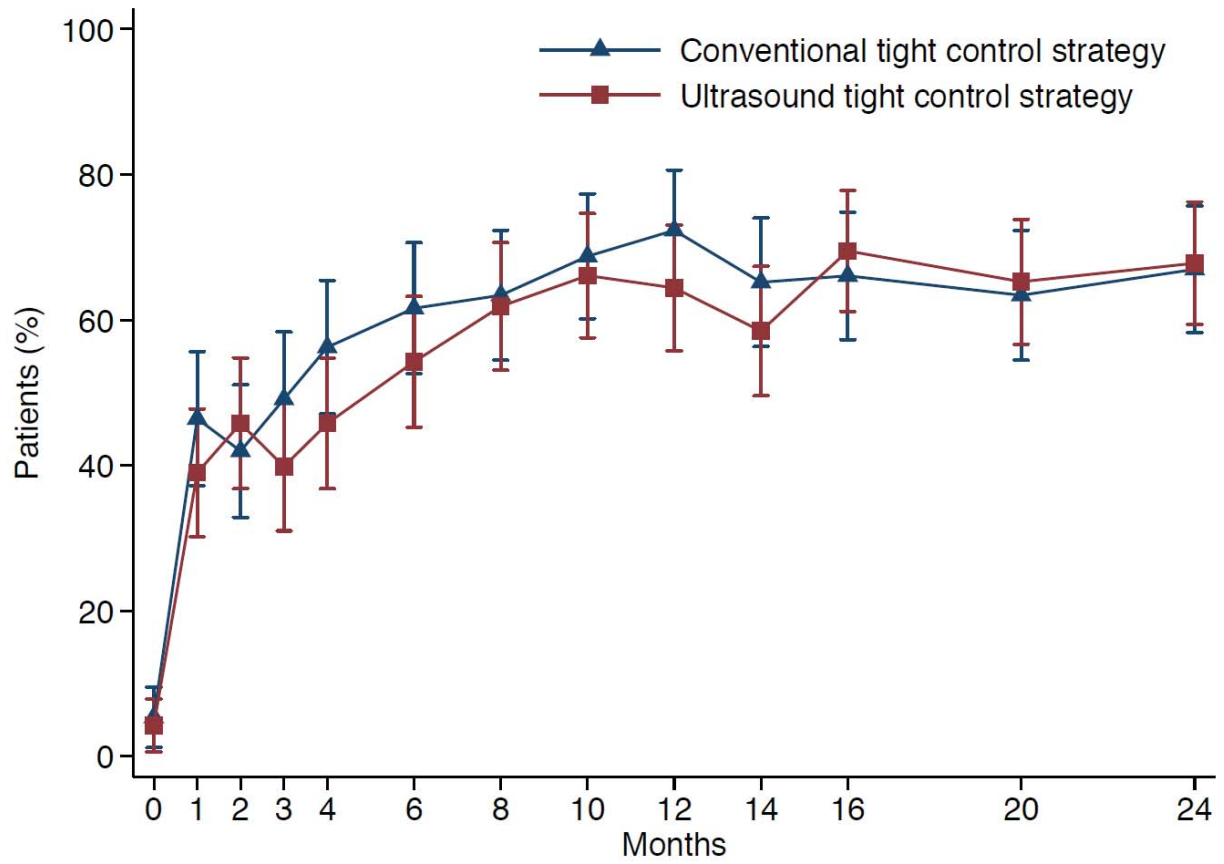


Figure 2C: Proportion of patients who achieved SDAI remission over 24 months derived from a logistic regression model. Bars indicate 95% confidence limits.

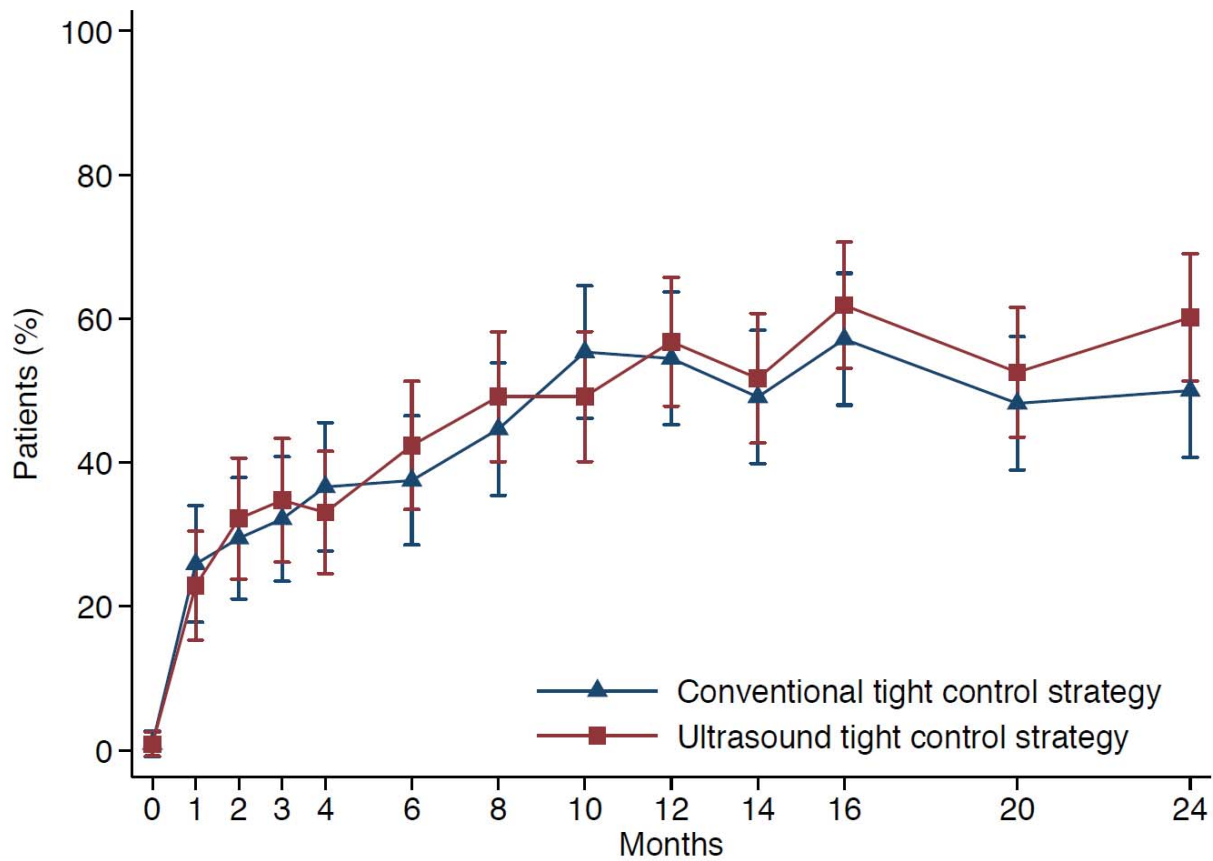


Figure 2D: Proportion of patients who achieved ACR/EULAR Boolean remission over 24 months from a logistic regression model. Bars indicate 95% confidence limits.

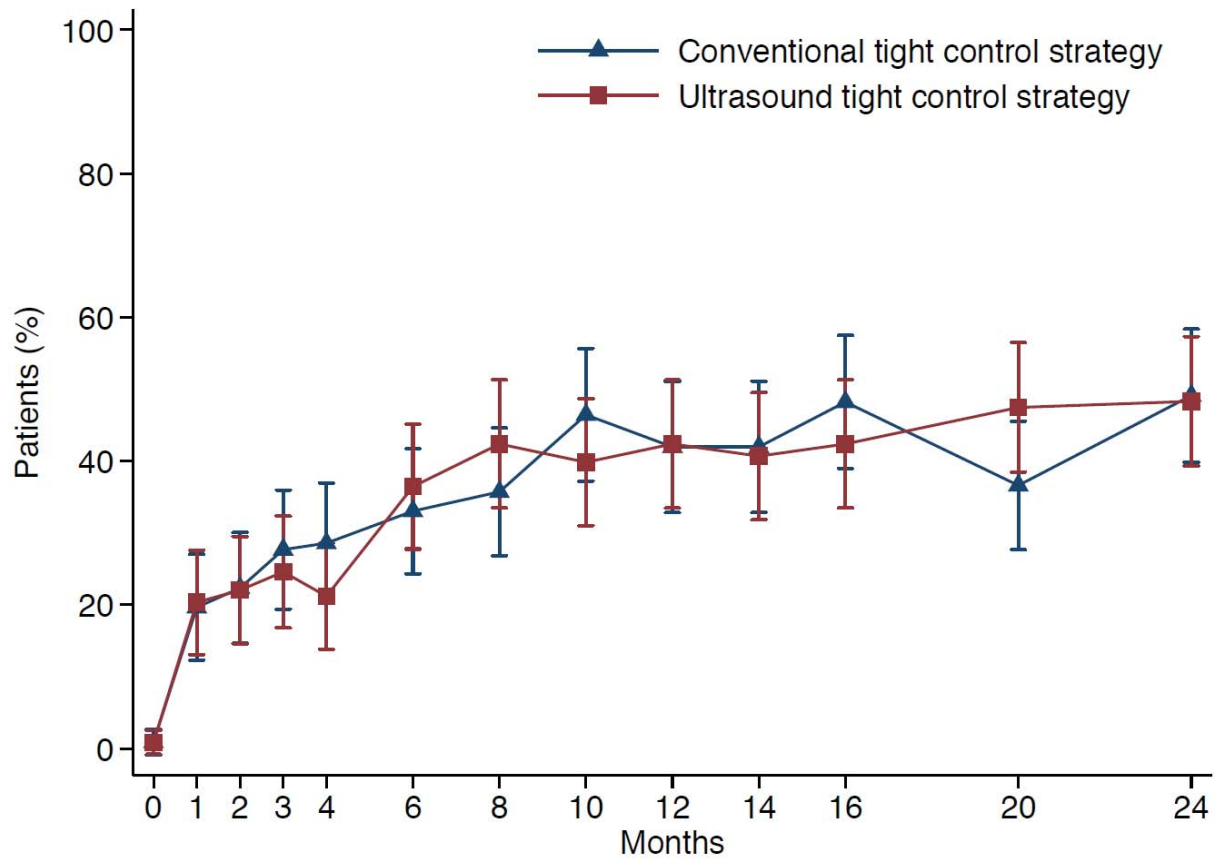
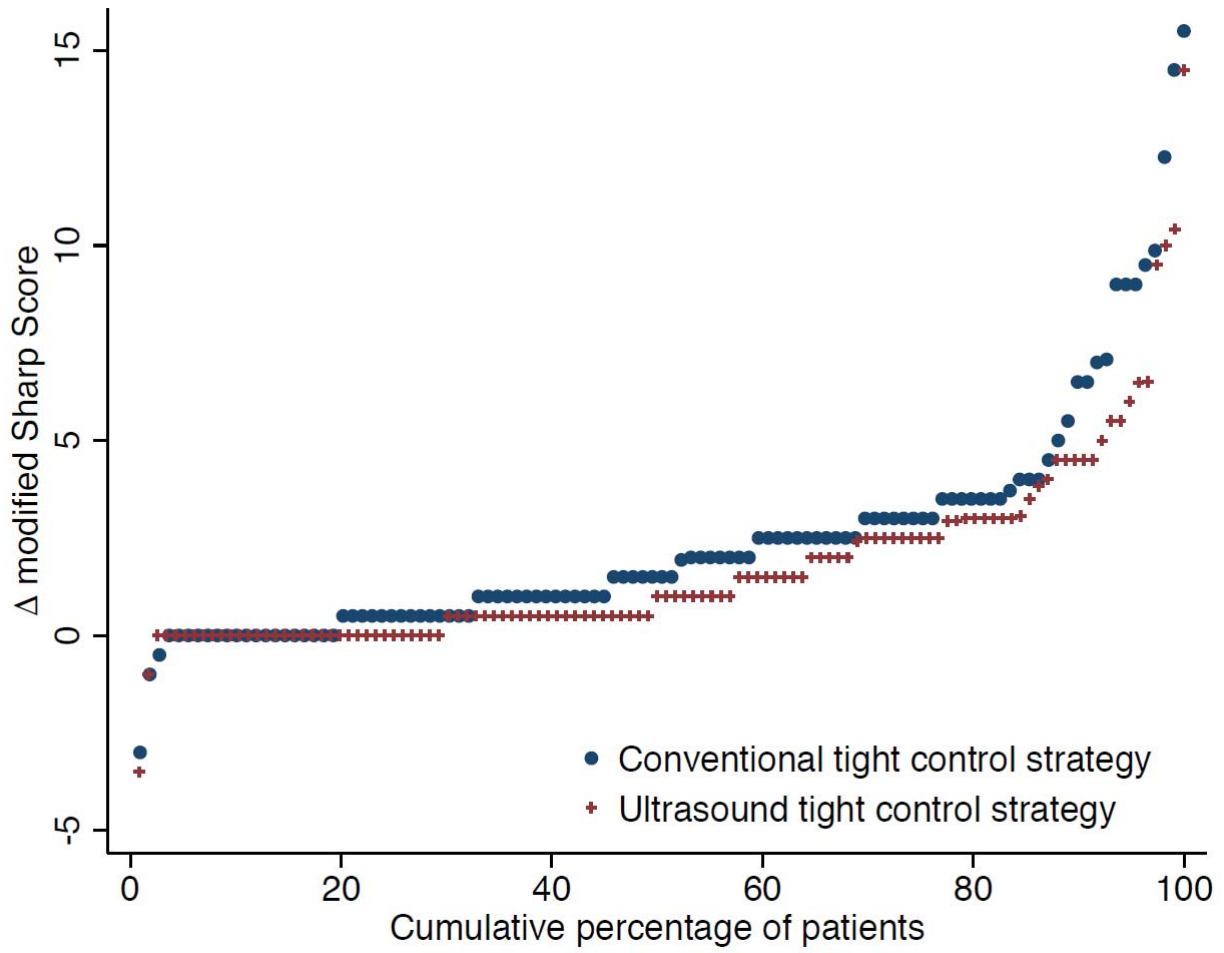


Figure 3. Cumulative probability plot of change between baseline and 24 months in van der Heijde modified Sharp score.



References

1. Grigor C, Capell H, Stirling A, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;**364**(9430):263-9.
2. Verstappen SM, Jacobs JW, van der Veen MJ, et al. Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. *Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial)*. *Ann Rheum Dis* 2007;**66**(11):1443-9.
3. Smolen JS, Aletaha D, Bijlsma JW, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;**69**(4):631-7.
4. Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet* 2009;**373**(9664):659-72.
5. Breedveld FC, Weisman MH, Kavanaugh AF, et al. The PREMIER study: A multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;**54**(1):26-37.
6. Genovese MC, Bathon JM, Martin RW, et al. Etanercept versus methotrexate in patients with early rheumatoid arthritis: two-year radiographic and clinical outcomes. *Arthritis Rheum* 2002;**46**(6):1443-50.
7. Weinblatt ME, Keystone EC, Furst DE, et al. Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum* 2003;**48**(1):35-45.
8. St Clair EW, van der Heijde DM, Smolen JS, et al. Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. *Arthritis Rheum* 2004;**50**(11):3432-43.
9. Aga AB, Lie E, Uhlig T, et al. Time trends in disease activity, response and remission rates in rheumatoid arthritis during the past decade: results from the NOR-DMARD study 2000-2010. *Ann Rheum Dis* 2015;**74**(2):381-8.
10. Smolen JS, Breedveld FC, Burmester GR, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2015.
11. Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;**73**(3):492-509.
12. Stoffer MA, Schoels MM, Smolen JS, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search update. *Ann Rheum Dis* 2015.
13. Singh JA, Saag KG, Bridges SL, Jr., et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Care Res (Hoboken)* 2015.
14. Lillegraven S, Prince FH, Shadick NA, et al. Remission and radiographic outcome in rheumatoid arthritis: application of the 2011 ACR/EULAR remission criteria in an observational cohort. *Ann Rheum Dis* 2012;**71**(5):681-6.
15. Felson DT, Smolen JS, Wells G, et al. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;**70**(3):404-13.
16. Szkudlarek M, Wakefield RJ, Backhaus M, et al. The discriminatory capacity of ultrasound in rheumatoid arthritis: active vs inactive, early vs advanced, and more. *Rheumatology (Oxford, England)* 2012;**51 Suppl 7**:vii6-9.
17. Dougados M, Jousse-Joulin S, Mistretta F, et al. Evaluation of several ultrasonography scoring systems for synovitis and comparison to clinical examination: results from a prospective multicentre study of rheumatoid arthritis. *Ann Rheum Dis* 2010;**69**(5):828-33.
18. Mandl P, Naredo E, Wakefield RJ, et al. A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. *The Journal of rheumatology* 2011;**38**(9):2055-62.
19. Saleem B, Brown AK, Quinn M, et al. Can flare be predicted in DMARD treated RA patients in remission, and is it important? A cohort study. *Ann Rheum Dis* 2012;**71**(8):1316-21.

20. Brown AK, Conaghan PG, Karim Z, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008;**58**(10):2958-67.
21. Brown AK, Quinn MA, Karim Z, et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. *Arthritis Rheum* 2006;**54**(12):3761-73.
22. Peluso G, Michelutti A, Bosello S, et al. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. *Ann Rheum Dis* 2011;**70**(1):172-5.
23. Scire CA, Montecucco C, Codullo V, et al. Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. *Rheumatology (Oxford, England)* 2009;**48**(9):1092-7.
24. Colebatch AN, Edwards CJ, Ostergaard M, et al. EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis. *Ann Rheum Dis* 2013;**72**(6):804-14.
25. Haavardsholm EA, Lie E, Lillegraven S. Should modern imaging be part of remission criteria in rheumatoid arthritis? *Best Pract Res Clin Rheumatol* 2012;**26**(6):767-85.
26. Wakefield RJ, D'Agostino MA, Naredo E, et al. After treat-to-target: can a targeted ultrasound initiative improve RA outcomes? *Postgrad Med J* 2012;**88**(1042):482-6.
27. Ostergaard M, Moller-Bisgaard S. Rheumatoid arthritis: Is imaging needed to define remission in rheumatoid arthritis? *Nature reviews Rheumatology* 2014;**10**(6):326-8.
28. van der Heijde D. Remission by imaging in rheumatoid arthritis: should this be the ultimate goal? *Ann Rheum Dis* 2012;**71 Suppl 2**:i89-92.
29. Moller-Bisgaard S, Horslev-Petersen K, Ejbjerg BJ, et al. Impact of a magnetic resonance imaging-guided treat-to-target strategy on disease activity and progression in patients with rheumatoid arthritis (the IMAGINE-RA trial): study protocol for a randomized controlled trial. *Trials* 2015;**16**:178.
30. Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;**69**(9):1580-8.
31. Hammer HB, Bolton-King P, Bakkeheim V, 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;**70**(11):1995-8.
32. Bruynesteyn K, Van Der Heijde D, Boers M, et al. Detecting radiological changes in rheumatoid arthritis that are considered important by clinical experts: influence of reading with or without known sequence. *The Journal of rheumatology* 2002;**29**(11):2306-12.
33. Hetland ML, Stengaard-Pedersen K, Junker P, et al. Aggressive combination therapy with intra-articular glucocorticoid injections and conventional disease-modifying anti-rheumatic drugs in early rheumatoid arthritis: second-year clinical and radiographic results from the CIMESTRA study. *Ann Rheum Dis* 2008;**67**(6):815-22.
34. Horslev-Petersen K, Hetland ML, Junker P, et al. Adalimumab added to a treat-to-target strategy with methotrexate and intra-articular triamcinolone in early rheumatoid arthritis increased remission rates, function and quality of life. The OPERA Study: an investigator-initiated, randomised, double-blind, parallel-group, placebo-controlled trial. *Ann Rheum Dis* 2014;**73**(4):654-61.
35. Detert J, Bastian H, Listing J, 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* 2013;**72**(6):844-50.
36. O'Dell JR, Mikuls TR, Taylor TH, et al. Therapies for active rheumatoid arthritis after methotrexate failure. *N Engl J Med* 2013;**369**(4):307-18.
37. Dale J, Stirling A, McInnes I, et al. Targeting Ultrasound Remission In Early Rheumatoid Arthritis – Results Of The Taser Study. *Arthritis & Rheumatism* 2013;**65 (Suppl)**.

Online only supplementary material:

**Ultrasound in the management of rheumatoid arthritis:
ARCTIC randomised controlled strategy trial**

Table of contents

Section 1. Study investigators in the ARCTIC trial	2
Section 2. Full inclusion and exclusion criteria	3
Section 3. Treatment regimen	4
Section 4. Statistical analysis	5
Section 5. Summary narratives for malignancies (n=5)	6
Section 6. Summary narratives for deaths (n=1)	6
Section 7. Statistical considerations: the conclusion of the ARCTIC study.....	7
Table S2. ARCTIC decision rules.....	9
Table S3. Analyses of Primary and Key Secondary Endpoints Adjusted for Sex.....	10
Table S4. Analyses of primary and key secondary endpoints, completer analysis set.....	11
Table S5. Distribution of women by centre and intervention.....	12
Table S6. Serious Adverse Events over 24 months (One Patient per Term).....	13
References web extra material	14
Section 8. Final protocol.	15
Section 9. Cover letter to statistical analysis plan.	61
Section 10. Statistical analysis plan.	62

Section 1. Study investigators in the ARCTIC trial

Study investigators in the ultrasound tight control arm:

Anna-Birgitte Aga; Hilde Berner Hammer; Inger Jorid Berg; Pernille Bolton-King; Åse Stavland Lexberg; Yngvill Hovde Bragnes; Hilde Stray; Merete Valen; Ellen Norli; Hilde Haukeland; Liv Elisabeth Kjustad; Shagaye Nabizadeh; Christian Høili; Per Jarle Tungevang; Ruth Stoklund Thomsen; Tina Pedersen; Vivi Bakkeheim; Gunnstein Bakland; Lucius Bader; Trude Jannecke Bruun; Hallvard Fremstad; Maud-Kristine Aga Ljoså; Anne Noraas Bendvold

Study investigators in the conventional tight control arm:

Till Uhlig; Liv Lefsaaker; Maria Karolina Jonsson; Tor Madland; Geirmund Myklebust; Inger Johanne Hansen; Anne Julsrud Haugen; Eva Purinszky; Erik Rødevand; Marit Seip; Synøve Kalstad; Carina Skorpen; Karen Irgens; Leif Kåre Haga; Halvard Dovland

Section 2. Full inclusion and exclusion criteria

Overview

Adult men and women with early RA according to the 2010 ACR/EULAR classification criteria with indication for DMARD therapy were eligible for participation in this study.

Inclusion criteria

To be eligible for this study, patients must meet all of the following criteria:

1. Male or non-pregnant, non-nursing female
2. > 18 years of age and < 75 years of age
3. Patients classified as having RA (according to new ACR/EULAR criteria)
4. Disease duration less than 2 years (defined as time from 1st joint swelling)
5. The treating rheumatologist decides the patient requires DMARD-treatment
6. The patient has taken no prior DMARD
7. Patients able and willing to give written informed consent and comply with the requirements of the study protocol

Exclusion criteria

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

1. Abnormal renal function (serum creatinine > 142 $\mu\text{mol/L}$ in female and > 168 $\mu\text{mol/L}$ in male, or GFR < 40 mL/min/1.73 m²).
2. Abnormal liver function (ASAT/ALAT > 3* normal), active or recent hepatitis, cirrhosis.
3. Major co-morbidities like severe malignancies, severe diabetic mellitus, severe infections, uncontrollable hypertension, severe cardiovascular disease (NYHA class 3-4) and/or severe respiratory diseases.
4. Leukopenia and/or thrombocytopenia.
5. Inadequate birth control conception, pregnancy, and/or breastfeeding.
6. Indications of active tuberculosis.
7. Psychiatric or mental disorders, alcohol abuse or other abuse of substances, language barriers or other factors which makes adherence to the study protocol impossible.

Section 3. Treatment regimen

Patients in both groups were treated according to the same fixed treatment algorithm, adhering to a treat-to-target strategy with DMARD escalation therapy if target was not met. The treatment adjustments (including i.a. injections) that could be made were defined in a pre-specified dosing regimen, outlined in table S1.

Treatment target

The treatment target in the ultrasound tight control strategy was clinical remission (defined as Disease Activity Score <1.6 and no swollen joints) and ultrasound imaging remission (defined as no power Doppler signal in any of the joints assessed by ultrasound). The treatment target in the conventional tight control strategy was clinical remission (Disease Activity Score <1.6 and no swollen joints). The ultrasound standardized score included assessments of the following 32 joints with both grey-scale and power Doppler (semi-quantitative score of 0-3 for all joints, with a reference atlas showing the different possible grades for all assessed joints): MCPs I-V, wrist (radio-carpal, radio-ulnar and inter-carpal), elbow, knee, talo-crural and MTP I-V bilaterally.[1]

Treatment adjustments

The decision of whether to adjust medication was based on change in and the level of the Disease Activity Score. If the patient does not respond as described in table S2, the treating physician immediately adjusted the therapy by proceeding to the next step in the treatment algorithm. If a patient responded or had reached the target, current medication was continued. In the ultrasound tight control group, the physician should overrule the decision based on the Disease Activity Score and proceed to the next step based on ultrasound findings, as described in table S2.

Intra-articular steroids

In both groups, clinically swollen joints were treated by intra-articular steroids when indicated. In the ultrasound tight control group an additional target was all joints with power Doppler signal, and all injections should be ultrasound guided. For both groups, intra-articular injections of only tender joints were not allowed. The maximum dosage of triamcinolone hexacetonid per visit was 80 mg which could be distributed within joints as decided by the treating rheumatologist.

NSAIDs, vitamin D and calcium

NSAIDs and coxibs were permitted. The choice and dosage of NSAIDs/coxibs was at the discretion of the treating rheumatologist. Analgesics up to the maximum recommended dose could be used for pain relief as required. Patients should avoid analgesics within 24 hours prior to a visit if possible.

All patients received vitamin D and calcium supplement during treatment with corticosteroids ≥ 7.5 mg, and postmenopausal women and older men (>70 year) was considered for a bisphosphonate according to general guidelines. IV or IM corticosteroids were not allowed during the study. Oral corticosteroids were allowed as described in table S1. Other DMARDs than those described in table S1 was not allowed.

Section 4. Statistical analysis

The full analysis set for efficacy and safety included all patients randomly assigned to a treatment group and who started the allocated intervention defined as having completed at least one regular visit after the baseline visit.

The primary analysis on the primary endpoint and other binary endpoints were conducted using logistic regression models. The analyses were not adjusted for the stratification factors center and presence of anti-CCP due to low cell frequencies, but these variables were included in robustness analyses using exact logistic regression. Estimates of risk difference were calculated from the logistic regression parameters using the delta method to provide the confidence intervals.

Missing values of the primary endpoint were imputed using the following rule:

Radiographic score:

- If the radiographic score was missing at month 24, the patient was considered not to meet the primary endpoint (worst outcome)
- If a radiographic score was missing for visit 11 (16 months), we used last radiographic observation

Disease Activity Score (DAS):

- If unable to calculate DAS at visit 13 (month 24), the patient was considered not to meet the primary endpoint (worst outcome)
- If unable to calculate DAS at visit 11 or 12 (month 16/20), we used last DAS observation

Swollen Joint Count 44 (SJC44):

- If SJC44 at visit 13 (month 24) was missing, the patient was considered not to meet the primary endpoint (worst outcome)
- If SJC44 was missing at visit 11 or 12 (month 16/20), we used last SJC44 observation

Other binary endpoints were imputed with worst outcome.

The radiographic scores by the van der Heijde modified Sharp method (total, erosion and joint space narrowing) change from baseline was analyzed using median regression with baseline value, center and presence of anti-CCP as covariates. Estimates of treatment difference and corresponding confidence intervals were computed using 10 000 bootstrap replications. Missing values were imputed using the following rule:

- Imputation by linear interpolation was used when observations existed both before and after the missing value
- Imputation by linear extrapolation using the last two known observations was used when no later observation existed

Binary variables derived from the radiographic scores (e.g. progression or not progression) was derived from the imputed data. The imputation method for radiographic scores was changed from the original statistical analysis plan. The original plan was to handle missing data for radiographic scores using multiple imputations, similar to other continuous endpoints. The change of method was done in order to conform with the typical analyses performed for radiographic scores.

Other continuous variables were analyzed using analysis of covariance adjusted for baseline value in addition to center and presence of anti-CCP. Missing values were handled using multiple imputations with 10 imputations drawn from the observed distribution using the Markov-chain Monte Carlo method.

Post-hoc robustness analyses were performed on the primary and a selection of secondary endpoints. The first robustness analyses addressed the skewed distribution of women between the treatment groups by adding sex as a covariate in the logistic and median regression analyses (table S3). The second robustness analyses addressed if the handling of missing data for non-completers affected the results, by restricting the analyses to completers only (table S4).

Descriptive statistics are presented using imputed values (worst outcome) for dichotomous endpoints and non-imputed values for continuous endpoints.

All analyses were done using Stata version 14.0 (StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.). All significance tests were two-sided, and we used 95% confidence limits. Secondary analyses were not adjusted for multiplicity.

Section 5. Summary narratives for malignancies (n=5)

1. A 66-year-old male in the conventional tight control group developed basal cell carcinoma an unknown date between four and 12 months after study start. During this period of time, she received triple synthetic DMARD therapy (methotrexate 20 mg weekly, salazopyrine 500-1000 mg twice daily, hydroxychloroquine 400 mg daily). Concomitant medication was folic acid. Medical history included unspecified cancer and osteoarthritis. The investigator considered the event not to be related to medications.

2. A 68-year-old male in the conventional tight control group was diagnosed with squamous cell carcinoma approximately 12 months after study start. Medical history included hypertension, diabetes, angina pectoris, myocardial infarction, cardiac surgery and lung disease. The patient was receiving 20 mg methotrexate weekly at the time of the event. Concomitant medications included metoprolol, lisinopril, atorvastatin, clopidogrel, acetylsalicylic acid, cetirizine, metformin, budesonide, mometasonefuroate, folic acid, vitamin B12 and B6. The patient was withdrawn from the study shortly after the occurrence of the event. He was treated with curative chemotherapy and radiation, and approximately four months after the occurrence of the event, the patient was reported recovered with no sequelae.

3. A 59-year-old female in the ultrasonography tight control group experienced a serious adverse event of cancer with liver metastases approximately two years after study start. The patient was receiving 20 mg methotrexate weekly at the time of the event. Concomitant medications included folic acid, colecalciferol and calcium. Medical history included pollen allergy, ankle fracture, periodical nausea, abdominal pain, elevated CRP and ESR. The patient was withdrawn from the study in response to the event and did not recover during the follow-up period. The investigator considered the event to be not related to the study drug.

4. A 49-year-old female in the conventional tight control group experienced a serious adverse event of breast cancer approximately four months after study start. Medical history included back pain. The patient was receiving 20 mg methotrexate weekly at the time of the event. Concomitant medications were folic acid, colecalciferol and calcium. The patient was withdrawn from the study five months after the onset of event and did not recover during the follow-up period. The investigator considered the event not likely related to the study drug.

5. A 70-year-old male in the ultrasonography tight control group experienced a serious adverse event of follicular lymphoma approximately two years after study start. Medical history included an unspecified type of cancer, cardiac disease, lung disease and arthrosis. The patient was receiving tiotropium, formoterol, acetylsalicylic acid, metoprolol, atorvastatin, folic acid, colecalciferol and calcium. Methotrexate was stopped approximately one year prior to the onset of the event. The patient had not recovered at the time of reporting.

Section 6. Summary narratives for deaths (n=1)

1. A 66-year-old male with a medical history of hypertension in the ultrasonography tight control group died from pneumocystis jirovecii pneumonia approximately 10 months after study start. Suspect medications were methotrexate 22.5 mg weekly and etanercept 50 mg weekly. Concomitant medications included lisinopril, alendronate, folic acid, colecalciferol and calcium. He started treatment with methotrexate in February 2012 and etanercept in August 2012. He was in very good shape at the clinical visit in October, and had no side effects of the drugs. From November, however, he developed dyspnoea and was hospitalized. Pneumocystis jirovecii pneumonia was proven by PCR and widespread pulmonary fibrosis by computerised tomogram. The infection was treated, but due to the pulmonary fibrosis, further treatment was terminated. He died in January 2013. The national competent authority considered the event to be possibly related to etanercept.

Section 7. Statistical considerations: the conclusion of the ARCTIC study

The ARCTIC study did not reject the primary null hypothesis of the trial: “There is no difference in the probability of achieving complete DAS remission after 24 months of treatment between the two treatment regimens (applying vs not applying ultrasonography)”. Failure to show an effect does not automatically imply a lack of effect of the intervention; the true effect might also be insufficiently large to be discovered by the trial. Negative trials can be divided into two categories: 1) True negative trials where the trial can rule out clinically important effects, and 2) Inconclusive trials where important clinical effects cannot be ruled out. In this section we discuss why we suggest that the ARCTIC trial is a true negative trial.

A central aspect in the assessment of a negative study is to consider the size of a clinically important potential effect. In our sample size calculations, we aimed to power the trial at 80% to detect a 20% difference between the interventions (Protocol). This was based on the remission rates in previous studies, in addition to discussions with clinicians and the study team regarding the effect size needed in order to introduce ultrasound in clinical practice. During the data collection in ARCTIC, two equivalence studies of biosimilar drugs in RA have been published, both with an equivalence margin of $\pm 15\%$. [2, 3] The setting in these studies were however very different, as the studies compared biological treatments, with response rates as the primary outcomes.

The estimated treatment difference of the primary endpoint was 3.3% with a 95% confidence interval of -7.1 to 13.7. The confidence interval is completely within both the $\pm 20\%$ and $\pm 15\%$ margin, ruling out a clinical important difference between the treatments according to both our estimate of an important clinical effect and the stricter definition used for assessment of biosimilar drugs. For the components of the primary endpoint, we can rule out a clinical important difference for the disease activity endpoints (no swollen joints and DAS remission) for both definitions, but the confidence interval for the difference in radiographic progression includes the $\pm 15\%$ margin.

If the study was to be repeated, the power to detect a 20% difference in the primary endpoint from 19% in the control group would have been 89%. The corresponding power with 15% difference would have been 68%. This further supports our conclusion.

Table S1. Treatment regimen in the ARCTIC trial

Visit (months)	Treatment if no response (if response continue treatment at present step, see table S2)
1 (0)	A. Monotherapy* + Prednisolone: 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) 2. Concomitant folic acid 5 mg/week (1mg 5/7 days or 5 mg x 1/week) 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 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)* Joint injections allowed as indicated according to treatment arm.
3 (2)	A. Optimize monotherapy* Increase Methotrexate to 25-30 mg/week <i>Or increase sulphasalazine/hydroxychloroquine/leflunomide dose</i>
4 (3)	A. Monitor start-up regimen (no changes in medication allowed unless due to AE)* Joint injections allowed as indicated according to treatment arm.
5 (4)	B. Triple combination therapy (or other combination therapy if MTX not tolerated):† 1. Add sulphasalazine, step up over 4 weeks to 500mg 2 x 2 and 2. Add hydroxychloroquine 200mg 1 x 2
6 (6)	B. Optimize triple combination therapy:‡ Add Prednisolone 7.5 mg 1 x 1
7 (8)	C. DMARD‡ and 1st biologic:§ 1. Highest tolerable dose MTX* and 2. Add 1 st biologic (according to current Norwegian guidelines) <i>*Or sulphasalazine/hydroxychloroquine/leflunomide if MTX not tolerated</i>
8 (10)	C. DMARD and 1st biologic: Adjust dose/interval of 1 st biologic
9 (12)	D. DMARD‡ and 2nd biologic: Switch to 2 nd biologic (according to current Norwegian guidelines)
10 (14)	D. DMARD‡ and 2nd biologic: Adjust dose/interval of 2 nd biologic
11 (16)	E. DMARD‡ and 3rd biologic: Switch to 3 rd biologic (according to current Norwegian guidelines)
12 (20)	E. Optimize DMARD and 3rd biologic plus prednisolone: Adjust dose/interval of 3 rd biologic and/or add prednisolone 7.5mg
13 (24)	F. Continue medication according to standard clinical care

* If MTX is not tolerated, switch to subcutaneous methotrexate, then continue according to scheme. In case of AE or not tolerated even in low dose subcutaneous, switch to sulphasalazine or hydroxychloroquine 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 oedema 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. tocilizumab).

§ 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).

Table S2. ARCTIC decision rules*

	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	
Ultrasound 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[‡]

* To be applied at all visits except visit 2 and visit 4.

[†] 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.

Table S3. Analyses of Primary and Key Secondary Endpoints Adjusted for Sex*

Variable	Ultrasound tight control (n=118)	Conventional tight control (n=112)	Difference (95% CI)	p-value
Primary endpoint – no. (%) †	26 (22.0)	21 (18.8)	2.9 (-7.7 to 13.6)	0.55
Components of primary endpoint				
No swollen joints – no. (%) ‡				
At 16, 20 and 24 months	62 (52.5)	61 (54.5)	-3.4 (-16.5 to 9.7)	0.61
DAS remission – no. (%) ‡				
At 16, 20 and 24 months	64 (54.2)	58 (51.8)	4.0 (-9.1 to 17.2)	0.55
No radiographic progression – no. (%) ‡				
Between 16-24 months	49 (41.5)	39 (34.8)	6.6 (-6.2 to 19.4)	0.32
Radiographic joint damage				
Δ Modified Sharp score at 24 months †	1 (0 to 2.5)	1.5 (0.5 to 3)	-0.43 (-0.92 to 0.07)	0.09

* All results were derived from the full analysis set, which included all randomised patients who underwent at least one visit after baseline. Median values are given with interquartile range (IQR). DAS=Disease Activity Score.

† The primary endpoint was the proportion of patients meeting all the 3 following criteria: 1) Sustained clinical remission, defined as DAS<1.6 at 16, 20 and 24 months 2) No swollen joints at 16, 20 and 24 months (44 Swollen Joint Count) and 3) No progression (<0.5 units) in van der Heijde-modified total Sharp Score between 16 and 24 months.

‡ Missing data before 24 months imputed using last observation carried forward, and missing data at 24 months imputed using worst outcome.

† Values are observed, unadjusted median values given with interquartile range (IQR). Treatment difference is derived from a median regression model. Missing data were imputed using linear intra- and extrapolation.

Table S4. Analyses of primary and key secondary endpoints, completer analysis set*

Variable	Ultrasound tight control (n=104)	Conventional tight control (n=100)	Difference (95% CI)	p-value
Primary endpoint – no. (%) †	26 (25.0)	21 (21.0)	4.0 (-7.5 to 15.5)	0.50
Components of primary endpoint				
No swollen joints – no. (%) ‡				
At 16, 20 and 24 months	62 (59.6)	61 (61.0)	-1.3 (-14.8 to 12.0)	0.84
DAS remission – no. (%) ‡				
At 16, 20 and 24 months	64 (61.5)	58 (58.0)	3.5 (-9.9 to 17.0)	0.52
No radiographic progression – no. (%) ‡				
Between 16-24 months	49 (47.1)	39 (39.0)	8.1 (-5.4 to 21.7)	0.24
Radiographic joint damage				
Δ Modified Sharp score at 24 months †	1 (0 to 2.5)	1.5 (0.5 to 3)	-0.45 (-0.86 to -0.39)	0.03

* All results were derived from the completer analysis set, which included all randomised patients who underwent at least one visit after baseline and who completed the study. Median values are given with interquartile range (IQR). DAS=Disease Activity Score.

† The primary endpoint was the proportion of patients meeting all the 3 following criteria: 1) Sustained clinical remission, defined as DAS<1.6 at 16, 20 and 24 months 2) No swollen joints at 16, 20 and 24 months (44 Swollen Joint Count) and 3) No progression (<0.5 units) in van der Heijde-modified total Sharp Score between 16 and 24 months.

‡ Missing data before 24 months imputed using last observation carried forward, and missing data at 24 months imputed using worst outcome.

† Values are observed, unadjusted median values given with interquartile range (IQR). Treatment difference is derived from a median regression model. Missing data were imputed using linear intra- and extrapolation.

Table S5. Gender distribution by centre and intervention

Proportion of female patients (%)		
Centre	Ultrasound tight control (n=118)	Conventional tight control (n=112)
1	29/37 (78.4%)	21/35 (60.0%)
2	5/9 (55.6%)	3/8 (37.5%)
3	6/7 (85.7%)	5/6 (83.3%)
4	9/14 (64.3%)	5/12 (41.7%)
5	2/4 (50.0%)	1/5 (20.0%)
6	7/9 (77.8%)	4/9 (44.4%)
7	5/7 (71.4%)	4/7 (57.1%)
8	9/9 (100.0%)	4/7 (57.1%)
9	4/5 (80.0%)	1/3 (33.3%)
10	5/11 (45.5%)	5/15 (33.3%)
11	3/6 (50.0%)	4/5 (80.0%)

Table S6. Serious Adverse Events over 24 months (One Patient per Term)

MedDRA System Organ Class	Ultrasound tight control (n=118)	Conventional tight control (n=112)
Gastrointestinal disorders		Volvulus*
Infections and infestations	Abscess, bacterial* Pneumonia* Pneumocystis jirovecii pneumonia‡	Localised infection* Abscess*
Musculoskeletal and connective tissue disorders		Arthralgia*
Neoplasms benign, malignant and unspecified (incl cysts and polyps)	Follicle centre lymphoma, follicular grade I, II, III † Metastases to liver*	Breast cancer*
Nervous system disorders		Syncope*
Renal and urinary disorders	Nephrolithiasis*	
Surgical and medical procedures	Percutaneous coronary intervention†	Hospitalisation‡

* The patient was receiving methotrexate.

† The patient did not receive any study medication, methotrexate was stopped one year prior to the diagnosis of the follicle centre lymphoma. The same patient received percutaneous coronary intervention four months after the baseline visit, and was then treated with methotrexate (this was prior to the diagnosis of lymphoma).

‡ The patient was receiving etanercept and methotrexate.

§ The patient did not receive any study medication (methotrexate was stopped approximately three months prior to the onset of the event).

References

1. 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. *AnnRheumDis*. 2011; 70(11):1995-1998.
2. Yoo DH, Hrycaj P, Miranda P, Ramiterre E, Piotrowski M, Shevchuk S, et al. A randomised, double-blind, parallel-group study to demonstrate equivalence in efficacy and safety of CT-P13 compared with innovator infliximab when coadministered with methotrexate in patients with active rheumatoid arthritis: the PLANETRA study. *Ann Rheum Dis*. 2013 Oct; 72(10):1613-1620.
3. Emery P, Vencovsky J, Sylwestrzak A, Leszczynski P, Porawska W, Baranauskaite A, et al. A phase III randomised, double-blind, parallel-group study comparing SB4 with etanercept reference product in patients with active rheumatoid arthritis despite methotrexate therapy. *Ann Rheum Dis*. 2015 Jul 6.

