# ECONOMIC EVALUATION OF INTERVENTIONS FOR INFLAMMATORY RHEUMATIC JOINT DISEASES

Thesis by

Maria Knoph Kvamme

2014

**Diakonhjemmet Hospital**

**Department of Rheumatology**

**Oslo, Norway**

**University of Oslo**

**Faculty of Medicine**

**Oslo, Norway**

*Series of dissertations submitted to the*
*Faculty of Medicine, University of Oslo*
*No. 1963*

ISBN 978-82-8264-906-3



Cover: Hanne Baadsgaard Utigard.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

# TABLE OF CONTENTS

# ACKNOWLEDGEMENT

This thesis is the result of a collaboration between the Department of Rheumatology, Diakonhjemmet Hospital and the Department of Health Management and Health Economics, University of Oslo. I am grateful for the financial support from the South-Eastern Norway Regional Health Authority and the opportunity to be employed at Diakonhjemmet Hospital and for the material support from the Department of Health Management and Health Economics.

I am interested in how we can make the best out of available resources in health care. When I was offered the opportunity to evaluate the cost-effectiveness of treatments for inflammatory rheumatic joint diseases, I accepted with joy. Professor Tore Kristian Kvien at Diakonhjemmet Hospital contributed with data from a comprehensive longitudinal observational study; the NOR-DMARD study and applied for financial support from the South-Eastern Norway Regional Health Authority and professor Ivar Sønbø Kristiansen at the Department of Health Management and Health Economics contributed with his expertise in economic evaluation in health care.

I would like to give special thanks to my principal supervisor professor Ivar Sønbø Kristiansen. Ivar was very engaged and encouraging from the first meeting and all the way through. He was goal-oriented and at the same time gave room for freedom.

I would also like to thank my co-supervisor professor Tore Kristian Kvien. The combination of scientific expertise within the field of rheumatology combined with positive and including characteristics was always very inspiring.

I am indebted to a great number of people who have assisted and supported me during the work with this thesis. I am grateful to Elisabeth Lie at Diakonhjemmet Hospital for sharing invaluable knowledge about the NOR-DMARD register as well as answering my questions in the field of rheumatology. I would also like to acknowledge professor Till Uhlig at Diakonhjemmet Hospital for assistance with the Oslo Rheumatoid Arthritis register and for generously contributing with his wide knowledge in rheumatology and Elisabeth Orskaug and Ingunn Friede Tvete at the Norwegian Computing Centre who contributed with statistical advice in the start of the construction of the NOrwegian Rheumatoid Arthritis (NORA) model.

At the Department for Health Management and Health Economics, I would like to acknowledge all colleagues and friends. Special thanks to Tron Anders Moger for statistical advice and Arna Desser and Anastasia Mokienko with whom I shared the office, for invaluable good companionship and generously sharing their good spirits.

I would also like to thank my mother, father, sister and brother for always being there for me. A special thanks to my mother-in-law Marianne for all the invaluable time and love to my children while I was working.

To Fredrik, who made this work possible, and to Philip, Anna and Astrid: this thesis is for you.

## FUNDING

The South-Eastern Norway Regional Health Authority funded this PhD project. The Department of Health Management and Health Economics provided office space and working environment. Additional financial funding was received from Marie and Else Mustad´s Legacy and the Research Council of Norway through HERO – Health Economic Research Programme at the University of Oslo. The pharmaceutical company Pfizer funded a health economics course including travel and accommodation.

## LIST OF PAPERS

I. Kvamme MK, Kristiansen IS, Lie E, Kvien TK. Identification of cutpoints for acceptable health status and important improvement in patient-reported outcomes, in rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis. J Rheumatol. 2010 Jan;37(1):26-31.

II. Kvamme MK, Lie E, Kvien TK, Kristiansen IS. Two-year direct and indirect costs for patients with inflammatory rheumatic joint diseases: data from real-life follow-up of patients in the NOR-DMARD registry. Rheumatology (Oxford). 2012 Sep;51(9):1618-27.

III. Kvamme MK, Lie E, Uhlig T, Moger TA, Kvien TK, Kristiansen IS. Cost-effectiveness of TNF inhibitors in combination with synthetic DMARDs versus synthetic DMARDs alone in patients with rheumatoid arthritis: a model study based on two longitudinal observational studies. Submitted to Rheumatology (Oxford). Status: revise and resubmit.

# ABBREVIATIONS

| | |
|---|---|
| ACPA | Anti-Citrullinated Protein Antibody |
| ACR | American College of Rheumatology |
| AS | Ankylosing Spondylitis |
| bDMARD | biologic Disease-Modifying AntiRheumatic Drug |
| CI | Confidence Interval |
| DAS28 | Disease Activity Score based on 28 joint counts |
| DMARD | Disease-Modifying AntiRheumatic Drug |
| DRG | Diagnosis Related Group |
| EMA | European Medicines Agency |
| EQ-5D | EuroQol-5 Dimensions |
| EULAR | European League Against Rheumatism |
| EVPI | Expected Value of Perfect Information |
| FCA | Friction Cost Approach |
| HAQ | Health Assessment Questionnaire |
| HCA | Human Capital Approach |
| HRQoL | Health-Related Quality of Life |
| ICER | Incremental Cost-Effectiveness Ratio |
| IQR | Inter-Quartile Range |
| LIS | Legemiddelinnkjøpssamarbeid (the Drug procurement cooperation) |
| LOS | Longitudinal Observational Study |
| MAU instruments | Multi-Attribute Utility instruments |
| MCII | Minimal Clinically Important Improvement |
| MHAQ | Modified Health Assessment Questionnaire |
| MTX | Methotrexate |
| NORA model | NOrwegian Rheumatoid Arthritis model |
| NOR-DMARD | NORwegian Disease-Modifying AntiRheumatic Drug study |
| NSAID | Non-Steroidal Anti-Inflammatory Drug |
| ORAR | Oslo Rheumatoid Arthritis Register |
| PASS | Patient Acceptable Symptom State |
| PsA | Psoriatic Arthritis |
| QALYs | Quality-Adjusted Life Years |
| RA | Rheumatoid Arthritis |
| RF | Rheumatoid Factor |
| ROC | Receiver Operating Characteristic |
| RCT | Randomized Controlled Trial |
| sDMARD | Synthetic Disease-Modifying AntiRheumatic Drug |
| SF-36 | The Medical Outcomes Study 36-item Short Form |
| SF-6D | Short Form-6 Dimensions |
| TNFi | Tumour Necrosis Factor inhibitor |
| VAS | Visual Analogue Scale |

## SUMMARY

The effectiveness of biologic disease-modifying antirheumatic drugs (bDMARDs) in the treatment of rheumatoid arthritis (RA) has been documented in a range of randomised controlled clinical trials (RCTs). The medical costs of using bDMARDs are substantially higher than those of synthetic DMARDs (sDMARDs). The European League Against Rheumatism (EULAR) emphasized the importance of evaluating costs in addition to treatment effect in the 2010, as well as in the 2013 updated recommendations, on the use of DMARDs in RA. Further, the Norwegian guidelines on priority in the health care sector include three criteria: severity of the health state, effectiveness of treatment and cost-effectiveness. The overarching aim of this thesis was to compare costs and effectiveness of bDMARDs versus sDMARDs for patients with RA.

The thesis consists of three studies. We conducted three studies because we wanted to examine measures of effectiveness, costs and cost-effectiveness of treatments for inflammatory rheumatic joint diseases. In the first study, we aimed at investigating two methods for letting patients evaluate the effectiveness of treatment, for better understanding of the patient perspective on outcome assessment. Both methodologies are dichotomous. The first asks if the patient is in an acceptable health state: Patient Acceptable Symptom State (PASS) (yes/no) and the second asks if the patient has experienced a Minimal Clinical Important Improvement (MCII) (yes/no). The instruments were investigated for use in economic evaluations and generally for use as measurements for evaluation of treatment effectiveness.

In the second study, we aimed at investigating the costs incurred by patients with inflammatory rheumatic joint diseases including RA, ankylosing spondylitis (AS) and psoriatic arthritis (PsA) in Norway.

The objectives of the third study were to estimate the incremental costs and health benefits of adding tumour necrosis factor inhibitor (TNFi) to sDMARDs in routine care for RA patients who were treated with sDMARDs. An additional objective was to compare the incremental quality-adjusted life years (QALYs) and incremental cost-effectiveness ratios (ICERs) when using the EuroQol-5 Dimensions (EQ-5D) versus the Short Form-6 Dimensions (SF-6D).

The three studies used observational data from the Norwegian Disease-Modifying Antirheumatic Drug register (NOR-DMARD). From 2000, NOR-DMARD recruited patients with inflammatory joint diseases when they started treatment with sDMARDs and/or bDMARDs. We used data from patients included up to February 2012, when NOR-DMARD comprised 7,675 patients. In addition, we used data up to 2001 from the Oslo Rheumatoid Arthritis Register (ORAR), established in 1994. Patients in the capital city Oslo with an RA diagnosis were included in this register since 1994.

The results from the first study revealed that PASS and MCII are not appropriate instruments for valuation of the effectiveness in economic evaluations. To use a dichotomised variable for valuation of the effectiveness demands that the variable change between 0 and 1 in the area where the effectiveness can be seen, in addition to being valid and reliable. PASS indicated effectiveness already at bad health states and MCII varied substantially with method used for the assessment. The estimated values for classifying a patient as being in an acceptable symptom state measured with EQ-5D (0.69) or SF-6D (0.65) were low (results for RA with the 80% specificity method). The results indicate that patients can report a satisfactory symptom state even when they have problems with regular daily activities and experience pain. Such a symptom state is worse than the treatment aim

for patients with inflammatory rheumatic joint diseases today. The estimated change values in EQ-5D and SF-6D for having experienced a MCII varied with method used for the assessment. The 80% specificity approach indicated relatively large values (0.1) compared to the 75$^{th}$ percentile approach, which resulted in zero in change value in health-related quality of life (HRQoL) to have experienced a clinically important improvement in patients with RA, assessed by the EQ-5D.

Health care costs as well as costs related to production losses were included for evaluation in the second study. bDMARD treatment entails considerable drug cost, but the total costs (including health care costs and production losses) decline during the first two years of treatment in both RA, PsA and AS. The total costs are similar across RA, PsA and AS and both health care costs and production losses seem to be high in Norway compared to other European countries for these diagnoses. The annual health care costs for RA patients on sDMARD treatment were approximately €3,400 (NOK 26,300) and for patients on bDMARD treatment the costs were €19,600 (NOK 152,600). The annual costs including production losses (HC approach) for RA patients on sDMARD treatment were approximately €32,200 (NOK 250,900) and for patients on bDMARD treatment the corresponding costs were €60,900 (NOK 475,400). Costs were somewhat lower for PsA patients and slightly higher for AS patients.

Previous studies have reported lower cost estimates than our study, except from recent studies performed in the US. Such differences are probably due to a high cost level overall in Norway. Further, our study was performed more recently than most other similar studies and we included the costs of bDMARDs, which is not the case in all previous studies since this patient group is relatively new. The time point is important for cost-of-illness studies in RA, PsA and AS because more intensive treatment is used today than previously.

By partly using the results from the first and second study, we could estimate the cost-effectiveness of TNFi-treatment of RA in Norway. In the third study, we developed a model; the NOrwegian Rheumatoid Arthritis (NORA) model to simulate the 10-year disease course and resource use in RA. The main challenge in the development of the NORA model was to find a comparable patient group for the traditional treatment strategy, since our patients were not randomized before choice of treatment. We found a group in ORAR that was similar to patients in the TNFi-strategy from NOR-DMARD and adjusted for population differences in HRQoL at start in the model. We thus chose to use patients from ORAR to estimate the treatment effect of sDMARDs. The subgroup of patients from ORAR that best matched the patient population from NOR-DMARD who started with a TNFi was in a somewhat better health state. Thus, we did not have identical patient groups and it cannot be excluded that this might have affected the results. Two main strengths with our model were that we had direct utility data and follow-up in the TNFi-strategy of up to 10 years. A novel approach in the modelling was using health states directly based on level of HRQoL. Previous models have used disease characteristics as the basis for health states and thereafter used regression estimates of HRQoL based on disease characteristics.

The cost-effectiveness results from the NORA model are in the range of previous study estimates of ICERs for bDMARDs versus sDMARDs in RA patients. A Norwegian willingness-to-pay level of approximately €67,300 per QALY was assumed in the base case. The choice of instrument for measuring HRQoL is of importance for the conclusion. The NORA model results indicate that TNFi-treatment, compared to sDMARD treatment, is cost-effective when accounting for production losses

with ICERs of €39,841 using EQ-5D and €60,227 using SF-6D. Excluding production losses, TNFi-treatment is cost-effective using EQ-5D (ICER €61,285), but not SF-6D (ICER €92,557).

## SAMMENDRAG (NORWEGIAN SUMMARY)

Effekten av biologiske sykdomsmodifiserende legemidler (DMARDs) ved behandling av revmatoid artritt (RA) har blitt dokumentert i flere randomiserte kontrollerte studier. Kostnadene ved å behandle med biologiske legemidler er imidlertid vesentlig høyere enn ved å bruke syntetiske DMARDs. Den Europeiske organisasjonen mot revmatiske sykdommer (EULAR), fremhevet betydningen av å vurdere kostnader i tillegg til behandlingseffekt i anbefalingene for bruk av sykdomsmodifiserende legemidler for RA både i 2010 og i den oppdaterte versjonen i 2013. De norske retningslinjene for prioritering i helsevesenet inkluderer tre kriterier: alvorlighet, effekt av behandlingen og kostnadseffektivitet. Hovedmålet i denne avhandlingen var å sammenligne kostnader og effekt ved bruk av biologiske versus syntetiske DMARDs hos pasienter med RA.

Avhandlingen består av tre studier. De tre studiene gjorde vi fordi vi ville undersøke måleinstrumenter for effekt, kostnader og kostnadseffektivitet av behandlinger for inflammatoriske revmatiske leddsykdommer. Målet i den første studien var å undersøke to metoder for å la pasienter vurdere effekten av den behandlingen de har fått for å bedre forstå pasientenes perspektiv ved effektevaluering. Begge metodene er dikotome inndelinger der den første spør om pasienten er i en akseptabel helsetilstand (PASS) (ja/nei) og den andre spør om pasienten har erfart en klinisk viktig forbedring (MCII) (ja/nei). Metodene ble vurdert både i forhold til bruk i økonomiske evalueringer og generelt for bruk som måleinstrumenter for effekt av behandlingsintervensjoner.

I den andre studien var målet å undersøke kostnader for pasienter med inflammatoriske revmatiske leddsykdommer inkludert RA, ankyloserende spondylitt (AS) og psoriasis artritt (PsA) i Norge.

Målene i den tredje studien var primært å estimere inkrementelle kostnader og helseeffekter av å legge til en TNF-hemmer i tillegg til behandling med syntetiske DMARDs hos RA pasienter i vanlig klinisk praksis. Mål nummer to var å sammenligne inkrementelle kvalitetsjusterte leveår (QALYs) og inkrementelle kostnadseffektivitetsbrøk (ICERs) ved bruk av EQ-5D versus SF-6D.

De tre studiene brukte observasjonsdata fra den norske DMARD studien (NOR-DMARD). Fra 2000 rekrutterte NOR-DMARD pasienter med inflammatoriske revmatiske leddsykdommer ved oppstart av behandling med syntetiske og/eller biologiske DMARDs. Vi brukte data fra pasienter inkludert til og med februar 2012, da NOR-DMARD omfattet 7 675 pasienter. I tillegg brukte vi data til og med 2001 fra Oslo RA register (ORAR), etablert i 1994. Pasienter i Oslo med en RA diagnose er inkludert i registeret.

Resultatene fra den første studien viste at PASS og MCII ikke er hensiktsmessige instrumenter for verdisettingen av nytten i økonomiske evalueringer. Det å bruke en dikotom indikator for å vurdere nytte krever at den veksler fra 0 til 1 i det området effekten skjer i tillegg til at den må være reliabel og valid. PASS indikerte nytte allerede ved svært dårlige helsetilstander og MCII varierte substansielt med hvilken metode som bruktes for å bestemme den. De estimerte verdiene for å klassifisere en pasient til å være i en tilfredsstillende helsetilstand målt i EQ-5D (0.69) og SF-6D (0.65) var lave (resultat for RA med 80% spesifisitets metoden). Resultatene indikerer at pasienter kan rapportere en tilfredsstillende helsetilstand når de har problemer med å gå, problemer med å utføre daglige aktiviteter og når de har smerte. En slik helsetilstand er lavere enn målsetningen med behandling for pasienter med inflammatoriske revmatiske leddsykdommer i dag. Estimerte endringsverdier i EQ-5D og SF-6D for å ha oppnådd en klinisk viktig forbedring varierte avhengig av hvilken metode som ble brukt for å bestemme dem. 80% spesifisitets metoden indikerte relativt store verdier (0.1)

sammenlignet med 75 persentil metoden, som viste 0 i endringsverdi i helserelatert livskvalitet for å ha oppnådd en klinisk viktig forbedring for pasienter med RA, evaluert med EQ-5D.

Den andre studien inkluderte kostnader knyttet til både helsetjenester og produksjonstap på grunn av sykdommen. Resultatene viste at kostnadene ved behandling med biologiske DMARDs er høye, men at de totale kostnadene (inkludert helsetjenestekostnader og produksjonstap) minker i løpet av de to første årene med behandling for samtlige analyserte diagnoser; RA, PsA og AS. De totale kostnadene er på omtrent samme nivå for RA, PsA og AS og både helsetjenestekostnader og produksjonstap ser ut til å være høye i Norge sammenlignet med andre Europeiske land for disse diagnosene. Årlige helsetjenestekostnader for RA pasienter ved behandling med syntetiske DMARDs var ca 26 300 kr mens for pasienter med biologisk DMARD behandling var kostnadene 152 600 kr. Medregnet produksjonstap (HC metoden) blir årlig kostnad for syntetiske DMARDs 250 900, mens biologisk får 475 400 kr.

Tidligere studier har rapportert lavere kostnader enn vår studie, unntatt nylige studier fra USA. Det skyldes sannsynligvis et generelt høyt kostnadsnivå i Norge, at vår studie er utført senere enn lignende studier og at vi inkluderte kostnader for biologiske DMARDs, noe som ikke er tilfelle i alle tidligere studier da pasientgruppen som bruker biologiske DMARDs er relativt ny. Tidspunkt er viktig ved kostnadsstudier av RA, AS og PsA fordi idag bruktes en mer intensiv behandling enn tidligere.

Ved å til dels bruke resultatene fra den første og andre studien kunde vi estimere kostnadseffektiviteten av TNF-hemmer behandling av RA i Norge. I den tredje studien utviklet vi en modell; «den NOrske Revmatoid Artritt (NORA) modellen» for å simulere 10 års sykdomsutvikling og ressursbruk ved RA. Den største utfordringen i utviklingen av modellen var å finne en sammenlignbar pasientgruppe for den tradisjonelle behandlingsstrategien; ettersom våre pasienter ikke var randomiserte før valg av behandling. Vi fant en gruppe i ORAR som lignet på pasientgruppen fra NOR-DMARD som startet med en TNF-hemmer og justerte for populasjonsforskjeller i helserelatert livskvalitet (HRQoL) ved start i modellen. Vi valgte å bruke pasienter fra ORAR for å estimere effekt av behandling med syntetiske DMARDs. Undergruppen i ORAR som mest lignet på pasientpopulasjonen fra NOR-DMARD var i en noe bedre helsetilstand. Vi hadde altså ikke identiske pasientgrupper og det kan ikke utelukkes at dette kan ha innvirket på resultatene. To viktige fordeler med vår modell var at vi hadde tilgang til direkte data på HRQoL og oppfølgingstid på opptil 10 år i TNF-hemmer strategien. Modellen baserte seg på å bruke HRQoL som grunnlag for å fordele pasientene til tilstand i modellen. Tidligere modeller har gått omveien om sykdomskjennetegn og deretter benyttet regresjonsestimat for HRQoL for disse kjennetegnene.

Kostnad-effekt resultatene fra NORA modellen er i samme størrelsesorden som funn fra tidligere studier av ICERs for biologiske versus syntetiske DMARDs for RA-pasienter. Antatt betalingsvilje for en QALY i Norge er rundt 500 000 kr i utgangspunktet. Valg av måleinstrumentet for HRQoL får betydning for konklusjonen. De modellerte resultatene indikerer at TNF-hemmer behandling, sammenlignet med behandling med syntetiske DMARDs, er kostnadseffektivt hvis produksjonstap er inkludert i kostnadene. Resultatene viste ICERs på 296 019 kr med EQ-5D og 447 488 kr med SF-6D. Med eksklusjon av produksjonstap er behandling med TNF-hemmere kostnadseffektivt hvis effekten beregnes med EQ-5D (ICER 455 351 kr) men ikke hvis den beregnes med SF-6D (ICER 687 697 kr).

# 1 INTRODUCTION

The inflammatory rheumatic joint diseases have severe health consequences. The patients can suffer from swelling, tenderness, pain and destruction of joints, causing disability and increased mortality (1-6). Traditionally, these diseases have been treated with disease-modifying antirheumatic drugs (DMARDs). In the last 15 years, a new type of medication has been increasingly used. These medications are called biologic DMARDs (bDMARDs). The effectiveness of bDMARDs in the treatment of rheumatoid arthritis (RA) has been documented in a range of randomised controlled clinical trials (RCTs) and systematic reviews (7-11).

This thesis investigates the inflammatory rheumatic joint diseases in terms of their costs to society and the cost-effectiveness of treatments specifically for RA in Norway. Current priority recommendations in the health care sector use three criteria: severity of the health state, effectiveness of treatment and finally cost-effectiveness. The combination of these three criteria are the basis for prioritizing in the Norwegian health care sector (12, 13). This thesis is concerned with the third criteria: cost-effectiveness. The theme is relevant because the costs of using bDMARDs are high and current treatment recommendations suggest using bDMARDs in RA-patients with active disease and inadequate response to methotrexate (MTX) and/or other conventional synthetic DMARDs (sDMARDs) (14). No cost-effectiveness/cost-utility evaluation of treatment with tumour necrosis factor inhibitor (TNFi) + MTX versus sDMARDs for RA-patients has been performed previously in Norway.

I want to make the thesis understandable and interesting both for researches with an economic background and for researches with a medical background. Therefore, both in the first sections of economics and health economics and in the sections of inflammatory rheumatic joint diseases I start by presenting basic knowledge of the respective topics. Readers who are well familiar with the concepts can thus read these quickly or jump to following sections.

The background of the thesis consists of three main parts; 1) economics, health economics and economic evaluation, 2) a presentation of three inflammatory rheumatic joint diseases and 3) the motivation for performing an economic evaluation of treatments for RA in Norway and a review of previous evaluations in the field. In the section of inflammatory rheumatic joint diseases, I give an overview of the main measures of disease activity and health status used in the diseases, since evaluation of health effects is debated and choice of measure has consequences for the cost-effectiveness of treatments.

Following the introduction, the overall objective of the project is presented as well as the different parts of the project. We started by exploring two general concepts of patient reported health and what their relationship was to multi-attribute utility (MAU) instruments. Subsequently, we investigated both health care costs and costs due to lost productivity for the inflammatory rheumatic joint diseases RA, psoriatic arthritis (PsA) and ankylosing spondylitis (AS). Previous research on costs for the specific diseases is presented directly after the description of the disease. The third part of the thesis aimed at performing an economic evaluation of biologic versus synthetic DMARDs for patients with RA in a Norwegian setting. Methods and data from these three parts are described and explained in the thesis, followed by a presentation of results and a discussion, divided in applied methods and a general discussion. The thesis is concluded by policy implications of findings and a look ahead to future research in the explored fields.

## 1.1 Definitions and theoretical foundations

### 1.1.1 Economics and health economics

The Nobel Prize laureate Paul Samuelson has defined the scientific discipline of economics as follows:

"Economics is the study of how people and society choose to employ scarce resources that could have alternative uses in order to produce various commodities and to distribute them for consumption, now or in the future, among various persons and groups in society" (15, p. 4). In other words, choices involve **trade-offs**. If we decide to spend more on one thing, we have less to spend on something else. Choices matter because resources are scarce. In making choices we respond to **incentives**, for example price changes. When we **exchange** with others, we increase the range of choices for both. In order to make good choices we need to have and utilize **information.** Finally, the choices we make determine the **distribution** of wealth and income in our society. These five concepts are central in understanding economics (16).

Folland, Goodman and Stano state, "Health economics is the study of how resources are allocated to and within the health economy. The production of health care and its distribution across populations fall within this definition" (17, p. 1). Health economics may be divided into two areas: Economic evaluation of health programs and incentives/financing. In economic evaluation, systematic evaluation of two or more alternatives ("strategies") is usually done to inform policy decisions. Governments, as well as individuals, face choices constrained by available resources. While economic theory assumes that individuals make choices based on implicit evaluation of costs and benefits of alternatives, policy makers need explicit quantification and valuation of cost and benefits of different alternatives. Policy makers must decide on spending in the health care sector versus other sectors, and make choices about which health care programmes to fund.

A range of methods has been developed to evaluate costs and benefits of public programmes. Cost-benefit analysis (CBA) was developed almost a century ago in that early forms of CBA were used by the US Army Corps of Engineers to evaluate flood control and similar water systems. The introduction of Medicare and Medicaid in 1965 created an interest among American economists to use CBA in health care. Methods and standards improved from 1981 when all new US federal regulations were to be evaluated by CBA. In a CBA, benefits and costs of programmes are measured in monetary terms. The challenges in representing life years or health improvements in dollar terms led to the development of new and alternative methods. The collection of CBA and the newer tools are commonly called economic evaluation (17).

### 1.1.2 Theoretical foundations in economic evaluation

The theoretical foundation of cost-benefit analysis lies in **welfare economics**. In CBA, the individual consumer should be the one to value programme outcomes in monetary terms (18). The main objective in welfare economics is to provide an ethical framework for making thoughtful statements about whether some states of the world are socially preferably to others. Welfare economists have developed a framework, called the neo-classical framework, which is built on four normative principles (19, 20);

1) The utility principle (*i.e.* individuals rationally maximize their welfare by ordering options and choosing the preferred option).

2) Individual sovereignty (*i.e.* individuals are themselves the best judges of what contributes most to their utility and how much that contribution is).
3) Consequentialism (*i.e.* utility is derived only from the outcomes of behaviour and processes rather than the processes themselves or intentions that led to the outcomes).
4) **"** Welfarism is the proposition that the "goodness" of any situation (*e.g.*, resource allocation) be judged solely on the basis of the utility levels attained by individuals in that situation " (21, p. 377).

" Taken together, these four tenets require that any policy be judged solely in terms of resulting utilities achieved by individuals, as assessed by individuals themselves " (21, p. 377).

The concept of utility has included different meanings during its history. In welfare economics, the usual way of interpreting the concept is that utility numbers are a representation of an individual's preference ordering over bundles of goods or states of the world. An individual moving to a preferred state of the world is the same as an individual having a higher level of utility. The individual utility is a function primarily of goods and services consumed by the individual himself although some welfarist economists include other sources of utility than goods and services. The narrowly defined consequential principle is not intrinsic to welfarist economics and some welfarists have tried to broaden the concept to include also processes and procedures. Finally, welfarism is probably the most characteristic tenet of the four principles in welfare economics. In welfarism, individual utility characterize all outcomes and "social welfare" is normally understood to be a function of individual utilities only (19).

The Pareto principle implies that social welfare increases only if the welfare of any member of society increases and that no others are worse off. In welfarism the initial distribution of wealth and income is taken as given (19). Often, in real-world projects there are both gainers and losers. Two approaches have been suggested to address this situation: the social welfare function and the compensation principle (22). In this text, I will not discuss the social welfare function. Kaldor and Hicks suggested the compensation principle. Kaldor stated that: "a project is desirable if, with the project, it is *hypothetically* possible to redistribute income so that everyone becomes better off than without the project" (22, p. 22). Hicks' criterion implied that it should not be possible for the losers to bribe the gainers not to undertake the project. Hicks stated that: "…a project, *i.e.* what can be labelled a move from state A to state B, is desirable if, in state A, it is impossible to redistribute income so that everyone is made as well off as in state B" (22, p. 23). The difference between the Pareto principle and the compensation principle is that the Pareto principle implies actual compensation in monetary terms, while the compensation principle only implies potential compensation. Hypothetical compensation allows for focus on the change in efficiency when a new policy is considered. A new policy is desirable if the revenues exceed the costs. A redistribution according to the Pareto principle should be possible but implementation is not required. Whether the redistribution should actually be carried out, is another but important question (22).

There is no clear, single theoretical foundation of cost-effectiveness and cost-utility analysis (CEA/CUA). The foundation has been referred to as the **decision-making approach**. The methodologies used in the application of the analyses reflect contributions from researches with different backgrounds beyond economics. It may consequently be claimed that CEA/CUA have

been developed as an applied technique for allocating resources. Operations research has also been suggested as the roots for CEA/CUA. Operations research can be considered a sub-discipline of mathematics and is concerned with solving problems of transferring theory to practical applications by means of advanced analytical methods. Other researchers have searched for the theoretical roots of CEA/CUA in welfare economics. Welfare economics represents a comprehensive framework that answers questions that arise from the societal perspective. The values implicit in welfare economics are not shared by all decision makers, even if analysts choose the societal perspective. Therefore alternative formulations of social goals regarding health and health care have been developed. This perspective has been called the "extra-welfarist" perspective (23). Put simply, the welfarist rooted in welfare economics assumes that the aim of the health care systems is to maximize utility, while the extra-welfarist assumes that the aim is to maximise health.

**Extra-welfarism** is a normative framework that does not include the restriction that "social welfare" is a function of individual utilities only. The extra-welfarist approach differs from the welfarist in four main ways:

1. It allows use of other outcomes than utility
2. It allows other sources of valuation than the affected individuals
3. It permits weighing of outcomes according to other principles than preference-based utilities
4. It allows comparison between individuals in several dimensions (19).

In extra-welfarism, individual utilities can be included in an evaluation, as in welfarist approaches. However, also quality of utility, equity weights, characteristics and capabilities might be included. In health economics, health has become seen as the central focus of evaluations. Health or health gain is pursued and appreciated for its own sake and not only because it yields utility. Instead of individual utilities, the evaluative space can include an assembly of individual characteristics such as health, some of which might be measured in a cost-utility analysis. A number of sources for valuation can be used under the extra-welfarist approach. It might be the affected individual, an expert, a representative sample of the general public or an authoritative decision-maker. The outcomes can be weighted and this is often considered important in order to allow for equity considerations. Weight can be assessed according to the characteristics of the people receiving the health benefits. It can be related to their age, wealth, need or initial health state. Finally, the framework allows for interpersonal comparability in outcomes. In welfarist economics, individual utilities are normally considered impossible to compare between individuals. The relevant comparable outcomes in the extra-welfarist approach are for example health and capabilities. The use of health measures such as quality-adjusted life years (QALYs) allow comparison of individuals within a health domain. It makes it possible for analysts to address questions from decision-makers such as "health optimisation" as a policy objective. Health optimisation can include improvement of average health as well as diminishing health inequalities (19).

## 1.2 Economic evaluation of health care interventions

### 1.2.1 A decision analytic framework

An economic evaluation can be included in a more comprehensive decision-analytic framework, as suggested by Hunink *et al* (24). In this section, only an introduction to this topic is presented. The

main aims with decision analysis in health care are to achieve improved communication about clinical controversies and achieve better decisions. A systematic approach to decision making under the name PROACTIVE has been suggested (24) (Table 1).

Table 1. The PROACTIVE approach to decision making

| | |
|---|---|
| **P**roblem | Define the problem. What exactly is the problem? |
| **R**eframe | Reframe from multiple perspectives. What is important from the perspectives of the patient, physician, department, hospital, payer and the public policy maker? |
| **O**bjective | Focus on the objective. What is the goal of an intervention? Why is this important? |
| **A**lternatives | Consider all relevant alternatives. Consider wait-and-see, intervention and obtaining information. Do I know all the reasonable alternatives? |
| **C**onsequences and chances | Model the consequences and estimate the chances. What events may occur over time? What are the chances? |
| **T**rade-offs | Identify and estimate the value trade-offs. What are the values and value trade-offs? What are the monetary costs? (*Necessary prerequisites for an economic evaluation*) |
| **I**ntegrate | Integrate the evidence and values. Can I quantitatively integrate the evidence and values or do I need a quantitative estimate of expected value? If there are uncertainties, what is the overall expected value of each alternative? (*Constructing a decision model for economic evaluation)* |
| **V**alue | Optimize the expected value. How do I optimize the decision? Can I combine the desirable and undesirable outcomes into one multi attribute outcome? (*Refining the decision model for economic evaluation)* |
| **E**xplore and evaluate | Explore the assumptions and evaluate uncertainty. Can I generalize the results to other patients? What if the population for which I am choosing a public health program is somewhat different? What if the estimates in my model are not quite accurate? Would plausible changes in any variable change the recommended action? (*Evaluate the results of the model)* |

Hunink M, Glasziou P, Siegel J, Weeks J, Pliskin J, Elstein A, et al. 2001

### 1.2.2 What is an economic evaluation and what are the main types?

Economic evaluation may be defined as “the comparative analysis of alternative courses of action in terms of both their costs and consequences” (18 p. 9). The definition includes two concepts that characterize economic analysis: 1) costs and consequences/inputs and outputs and 2) choice of alternatives. These two concepts can be used for categorizing economic evaluation as one of several evaluation situations in health care (table 2).

Table 2. Distinguishing characteristics of health care evaluation

<table>
<tr><td></td><td colspan="4">Are both costs and consequences of the alternatives examined?</td></tr>
<tr><td rowspan="6">Is there comparison of two or more alternatives?</td><td></td><td colspan="2">No</td><td>Yes</td></tr>
<tr><td rowspan="3">No</td><td>Examines only consequences</td><td>Examines only costs</td><td></td></tr>
<tr><td colspan="2">Partial evaluation</td><td>Partial evaluation</td></tr>
<tr><td>1A Outcome description</td><td>1B Cost description</td><td>2 Cost-outcome description</td></tr>
<tr><td rowspan="2">Yes</td><td colspan="2">Partial evaluation</td><td>Full economic evaluation</td></tr>
<tr><td>3A Efficacy or effectiveness evaluation</td><td>3B Cost analysis</td><td>4 Cost-benefit<br>Cost-effectiveness<br>Cost-utility analyses</td></tr>
</table>

Drummond MF, Sculpher MJ, Stoddart GL, O'Brien BJ, Torrance GW. 2005

In 1A, 1B and 2 there is only a description of a programme or service and no comparison to another alternative. In 1A, the consequences of a programme is described, which is termed an outcome description. In 1B, when only costs are described, the result is called a cost description. These descriptions include cost of illness and burden of illness studies, which present the societal costs of a disease. In 2, both consequences and costs are described and these reports are labelled cost-outcome descriptions. Moving down in the table, 3A describes a situation in which two or more alternatives are compared but only consequences are compared and similarly in 3B only costs between alternative interventions are compared. All of these types of evaluations are called partial evaluations. Finally, in 4, different types of full economic evaluations are listed. The shared feature of the economic evaluations is that they include both an assessment of costs *and* consequences in addition to making a comparison of alternative interventions (18).

The first type, called cost-benefit analysis (CBA) measures both the costs and consequences of programmes in monetary terms. The results may be stated either as a sum (which can be negative) representing the net benefit (loss) of one alternative course of action versus another or as a ratio of costs to benefits. In order to allow for comparison of outcomes, analysts often attempt to go beyond the disease specific effects and attach a generic measure of value to an effect/a set of effects generated from an intervention. The consequences of an intervention are expressed in monetary terms and directly compared to the costs of the same programme. The challenge lies in the task of translating health effects, such as life-years gained, disability days avoided or QALYs gained into their corresponding monetary value. Most people have no experience with valuing benefits in monetary terms. In practise, the analysis is often restricted to the benefits and costs that can easily be expressed in monetary terms (18).

The second type, cost-effectiveness analysis (CEA), has been described in broad terms as: “a method used to evaluate the outcomes and costs of interventions designed to improve health” (23, p. 3). This description includes the prerequisites: costs *and* consequences as well as interventions, indicating that more than one option should be evaluated. Analyses in which costs are related to a single, common effect that may differ in size between alternative programmes are included in the term. The outcome may be life-years gained, disability-days avoided or a more specific effect such as cases of deep vein thrombosis detected or episode-free days in asthma. A CEA can be performed on any alternatives with a common effect (18). The results of a CEA are normally summarized in cost-effectiveness ratios that illustrate the cost of achieving one unit of health outcome, for example the

cost per life-year gained, for different interventions and patients. CEA presents explicitly the trade-offs when choosing among interventions by providing estimates of outcomes and costs. Interventions under comparison can be ranked on the basis of their cost-effectiveness ratios and the interventions with the lowest cost per year or case are the most efficient ways of improving health (23). The comparison of costs per output is based on the ratio of incremental costs to incremental effects. The change in costs using a new programme versus a standard programme can be described as $C_1$-$C_0$. Similarly, the change in health effects of using the new programme can be described as $E_1$-$E_0$. The two programmes can thus be compared using the incremental cost-effectiveness ratio (ICER):

$$\frac{C1-C0}{E1-E0} \quad (17)$$

C = Costs, E = Effectiveness

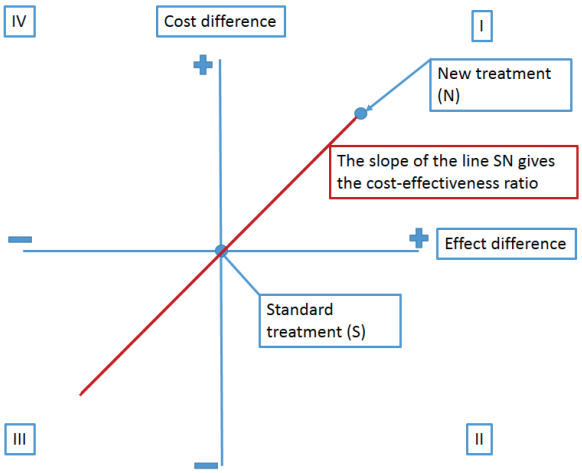
Finally, cost-utility analysis (CUA) encompasses evaluations that use utilities as the outcome measure. CUA is a variant of CEA, but it is a more comprehensive type of analysis than CEA because it includes valuation of outcomes. CUA is typically expressed as the cost per healthy year or the cost per QALY gained by implementing one intervention instead of another. The utility refers to the preferences individuals or the society have for a particular health state. The utility of a health state, outcome or effect can thus be different from the health state, outcome or effect in itself. For a given set of health outcomes of an intervention, health-related quality of life adjustment can be made. The resulting generic outcome measure can allow for comparison of costs and consequences in different interventions and in different patient groups. Generic outcome measures include QALYs, healthy years' equivalent (HYE), the disability-adjusted life year (DALY) and the saved-young-life equivalent (18). In a cost-utility analysis, the ICER is similarly expressed as:

$$\frac{C1-C0}{U1-U0} \quad (18)$$

C = Costs, U = Utilities

The ICER, or ICERs when more than two alternative programmes/interventions are compared, can be presented visually in the cost-effectiveness plane (figure 1). The cost-effectiveness plane has two dimensions: the y-axis, which presents the difference in costs and the x-axis, which presents the difference in effectiveness between a new intervention (N) and standard treatment (S). The relevant alternative to a new treatment could be status quo or a competing programme. The slope of the line SN gives the cost-effectiveness ratio. If the new intervention's point estimate lies in quadrant II, the new intervention is both more effective and less costly than S and it should clearly be implemented. On the contrary, if the point estimate of N is in quadrant IV, it is both less effective and more costly than S and should not be implemented. However, when the estimate of N lies in quadrants I or III, the implementation of the intervention depends on the maximum cost-effectiveness ratio which the decision maker is willing to accept (18), the maximum willingness-to-pay (WTP).

Figure 1. The cost-effectiveness plane (adapted from Drummond MF, Sculpher MJ, Stoddart GL, O'Brien BJ, Torrance GW (18).




### 1.2.3 Costs and cost analyses

When performing a cost analysis two overarching points should be considered:

First, the range of relevant cost items to include should be assessed. Four main categories of costs can be identified for health care programmes or treatments. These consist of:

1. Resources consumed in the health sector
2. Resources consumed in other sectors
3. Patient/family resources
4. Production losses/gains

The first category includes resources such as hospital stays, visits to general practitioners (GPs), physical therapists, private specialists, and rehabilitation stays etc. It also includes costs of medication and the costs of blood analyses. The second category includes costs incurred from other public agencies and the voluntary sector. However, this category is seldom taking into account in economic evaluation and may be insignificant in many interventions. The third category includes patient and family resources that they devote to the treatment process and out-of pocket expenses. The fourth category includes costs due to loss of time at work and lost productivity while at work (18).

The perspective of the analysis decides which cost categories to include. It can for example be the patients' perspective, the perspective of the health care provider or a societal perspective (18). In Norway, the societal perspective is recommended to use in economic evaluations in the health care sector (25). The societal perspective is also recommended in the literature as the point of view that should be adopted when in doubt, as it is the broadest perspective and is usually the most relevant approach (18).

Second, individual cost items should be identified and valued. This includes two main elements:

- Quantification of resource use
- Valuation of the resources (assessment of unit prices)

Measurement of quantities of resource use is often determined by the context of the study. The main data sources are clinical trials and registers. Data from clinical trials can be collected directly or they can come from published results.

The theoretically correct cost for a resource is its opportunity cost. This means "…the value of the foregone benefits because the resource is not available for its best alternative use" (18, p. 57). However, the opportunity cost is in many cases not identifiable. When there is perfect competition, the market price reflects the opportunity cost. When there is not, to use market prices is a pragmatic approach for the assessment of prices (18). Norwegian guidelines recommend using market prices for the estimation of unit prices. If there is no or little competition with private providers, the recommendation is to use prices exclusive of value-added tax (25). When market prices do not exist, other sources for valuation can be national fee schedules and diagnosis related group (DRG) price lists.

The value of leisure time has been debated in the literature and some argue that patients' time costs in receiving health care should be included if the economic evaluation is undertaken from a societal perspective (23). The most common approach is not to attach a monetary value of leisure time in the base case analysis. If it is included, patient's and family members' time may be valued by means of wage rates (18).

Both the questions of *whether* production losses should be included and *how* they should be included in an economic evaluation have been subjects to extensive debates in the literature (18, 23, 26). These issues relate both to the first point considered above: the range of relevant cost items to include and to the second: how the cost item should be valued. In this text, I will introduce the topic but I will not go into details in the debate.

First, should production losses be included in an economic evaluation? An argument against including production losses is that inclusion may entail double counting of the value of production gains if these are included in the denominator, *i.e.* the measure of effectiveness. When QALYs are used as the measure of health benefit, there might be double counting of production gains related to an intervention. Respondents might include what the value of return to work would have on their income when assessing their HRQoL (18). The Norwegian guidelines include this consideration: "production losses should in principle be included, but there is a question about possible double counting of production losses. This could be the case if these are partly included in the effectiveness measure when QALYs are used" (25, p. 21). Another concern is related to equity considerations. It has been argued that only production losses relating to resources that in alternative use could be used in health care are relevant to include (26). The Norwegian guidelines preliminary recommendation is that production losses can be included in the analysis if they are documented or can be substantiated. The analysis should present results both including and excluding production losses (25).

Second, how should production losses be valued in an economic evaluation? The typical way of valuing production losses are to use gross earnings, including employment overheads and benefits (18). The viewpoint in the Norwegian guidelines is to value production losses by using the mean income inclusive of tax, employment fees and other societal costs, specifically for the type of competence included in the intervention (25). It has been argued that using a general wage rate to value production losses rather than the actual wages of individuals affected by the intervention, could be one way of counteracting the equity consideration. Two general costing approaches have been recommended: the human capital approach (HCA) and the friction cost approach (FCA). In the HCA, the gross earning (including employment overheads and benefits) of those in employment are included in the estimate (18). The HCA has no time limit for how long a work loss should be accounted for. In contrast, in the FCA, production losses are only accounted for in a limited period, the "friction period", which is the assumed time it would take for an absent worker to be replaced by another. This period can vary both in types of work and in different parts of the work. The FCA gives lower estimates of production losses than the HCA (27).

#### *1.2.3.1 Discounting*

Effects and costs obtained in different years have to be adjusted to present values to allow for summary and comparison of different interventions. The recommended yearly rate of discounting, both for health effects and costs is currently four percent in Norway (25). The debate on whether or not costs and effects should be equally discounted will not be discussed in this text (28).

### 1.2.4 Uncertainty

Variability, heterogeneity and uncertainty are present in all economic evaluations (table 3). Uncertainty is a core factor in economic evaluation for decision making (29). Uncertainty analysis can have two main purposes: assess confidence in a chosen course of action and assess the value of seeking additional information to improve the decision information (30).

Table 3. Key concepts in uncertainty and heterogeneity in models for cost-effectiveness/cost-utility analysis

| Concept | Explanation |
|---|---|
| Variability/Stochastic uncertainty | Differences by randomness in outcomes between identical patients. Variability cannot be reduced by additional data. |
| Parameter uncertainty | The precision with which an input parameter is estimated. The parameter can for example be a mean cost, a mean utility or the probability of an event. Parameter uncertainty can be reduced by additional data collection. |
| Heterogeneity | The variability between patients who can be attributed to characteristics of those patients. For example, a specific event can be more likely in women over 70 years old. |
| Structural uncertainty | The assumptions inherent in the decision model. |

Briggs A, Claxton K, Schulper M (29) and Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD (30).

**Variability** between subjects is the differences in outcome that occur between patients by chance (29). Other terms used for this type of uncertainty are first order uncertainty, stochastic uncertainty

and Monte Carlo error (30, 31). An example of variability can be a group of 10 patients that will have a hip-replacement surgery. We consider these patients to be a homogenous group. Two patients will have an infection during a two-week period following the procedure, which corresponds to an estimated risk of 0.2 or 20%. Each individual has thus a 20% chance of having the infection. Since each individual will either have an infection or not, there will be a variability between patients even if the true probability in this case is 20% and there is no uncertainty about the probability of infection (29).

**Parameter uncertainty** is intrinsic to a cost-effectiveness/cost-utility model. It refers to the uncertainty in estimated parameter values, for example the risk of infection after hip surgery. Another name for this type of uncertainty is second order uncertainty. In the example above, the estimated probability of an infection was 20%, but this estimate may be uncertain. The standard approach to express the uncertainty is using a binomial distribution and estimate the standard error (SE) of the mean as follows:

SE ($p$) = $\sqrt{\mathrm{p}(1-p)/n}$ (29 p. 82)

$p$ = the estimated proportion (0.2)

$n$ = the sample size (10)

In the example above SE ($p$) = 0.13

The 95% confidence interval (CI) is obtained by taking 1.96*SE at each sides of the point estimate of 0.2. This results in a 95% CI of (-0.05,0.45).

If we had observed 20 infections in 100 surgical procedures, $p$ had still been 0.2, but the uncertainty of the estimate would have been much reduced *i.e*. SE ($p$) = 0.04 and the 95% CI would have been (0.12-0.28).

The distinction between variability and parameter uncertainty is illustrated in this example. Since the probability of 0.2 is unchanged, the variability between patients is the same. In contrast, the parameter uncertainty in the estimate is reduced from a SE = 0.13 to a SE = 0.04 and the associated 95% CI is reduced from (-0.05,0.45) to (0.12,0.28) around the point estimate of 0.2 (29).

All models include parameters that have to be estimated. Standard statistical methods for estimation generate a point estimate and a measure of precision, which can be a standard error or a 95% confidence interval. Primary data sources or one or more secondary sources can be used to inform input parameters. Irrespective of data source, the estimation generates a point estimate and a measure of precision. The estimation information should be used directly into the uncertainty analysis.

The representation of uncertainty depends on the type of uncertainty analysis. Parameter uncertainty can be assessed by deterministic sensitivity analysis or by probabilistic sensitivity analysis. In deterministic sensitivity analysis, an interval estimate representing the beliefs about the parameter's plausible range is necessary. In probabilistic sensitivity analysis, distributions are specified through its parameters. In a deterministic sensitivity analysis, the parameter values are

varied with the aim of testing the sensitivity of the results to a specific parameter or specific parameters. In probabilistic sensitivity analysis, all included parameters are usually varied at the same time. The parameter uncertainty is expressed by means of probability distributions. The parameter values are sampled from pre-defined probability distributions (30).

In the choice of distribution, the general principle is that assumptions for specifying the distribution should follow standard statistical methods. These include: Beta distributions for binomial data, Dirichlet distributions for multinomial data, gamma or log normal for right skewed data (for example cost data), log normal for relative risks or hazard ratios and logistic for odds ratios (29, 30).

The presentation of distributions below is limited to the distributions we used in the third study included in this thesis. The Dirichlet distribution is the multivariate generalization of the beta distribution. The beta distribution is constrained on the interval 0-1 and includes two parameters, α and β. The fitting of the distribution is illustrated by the following example: The number of events of interests are *r*, from a given sample size *n*. The point estimate of the probability is given as the proportion of events to the total sample. The uncertainty in the probability is simply given by setting α=r and β=n-r. In the Dirichlet distribution, the number of parameters are the same as the number of categories (*i.e.* health states) in the multinomial distribution (29).

Beta $(\alpha, \beta)$, $\alpha, \beta > 0$ logical constraints $(0 \leq \pi \leq 1)$

Dirichlet $(\alpha_1 \ldots \alpha_k)$, $\alpha_k > 0 \; \forall_k$ logical constraints $(0 \leq \pi \leq 1)$

Cost data are logically non-negative and consist of counts of resource use weighted by unit costs. The Poisson distribution, which is discrete, is often the candidate distribution for count data in standard statistics. The gamma distribution is the conjugate to the Poisson and is often used for cost parameters. The gamma is constrained on the interval 0 to positive infinity. To fit a gamma distribution the observed sample mean and variance are set equal to the corresponding expressions for mean and variance of the distribution:

$\overline{x} = \alpha\beta$, $s^2 = \alpha\beta^2$

$\alpha = \overline{x}^2/s^2$, $\beta = s^2/\overline{x}$

Gamma $(\alpha, \beta)$, $\alpha, \beta > 0$ logical constraints $(\theta \leq 0)$

Utility parameters are theoretically constrained on infinity at the lower end and one at the upper end, representing the worst possible health state and perfect health, respectively. To fit a gamma distribution a constraint on the distribution from 0 to positive infinity is necessary. We can transform the utility to decrements to fit the distribution: D=1-U. D represents the utility decrement/disutility and U represents utility (29).

International consensus has recommended reporting both deterministic sensitivity analysis and probabilistic sensitivity analysis. Deterministic sensitivity analysis can report variation in outcomes resulting from varying key parameters and probabilistic sensitivity analysis can report on overall uncertainty. In the reporting of a probabilistic sensitivity analysis, the specific distributions with its parameters should be presented in addition to a justification of the choice of distribution (30).

The results from a probabilistic sensitivity analysis can be represented in scatter plots in the cost-effectiveness plane and in cost-effectiveness acceptability curves (CEACs) and value-of-information analyses. A scatter plot in the cost-effectiveness plane shows the joint distribution of costs and outcomes from Monte Carlo simulations and gives a visual impression of the decision uncertainty. Cost-effectiveness acceptability curves show the probability that a strategy or an alternative is cost-effective given different values for a unit of health benefit. To better inform a decision, the value of additional research can be assessed in value-of-information analyses. These analyses are relevant when the decision maker has the authority to commission or mandate future research. Value-of-information analysis include the expected value of perfect information (EVPI) which is estimated for the total economic evaluation and the expected value of partial perfect information (EVPPI) which can be estimated for a specific parameter/specific parameters in the evaluation. EVPI and EVPPI require a probabilistic model. The EVPI combines the probability of making the wrong decision with the consequential cost of making the wrong decision. The higher the EVPI, the higher is the opportunity cost of making the wrong decision at the specific point at which the decision is being made. The EVPI is often reported in monetary terms, using net monetary benefit, but it can also be reported using net health benefit. The net monetary, as well as the net health benefit, depend on the ICER threshold, *i.e.* the willingness-to-pay. Consequently, the EVPI should be reported for a specified ICER threshold(s) (30). The EVPPI may be very computationally intensive.

**Heterogeneity** is the extent to which there is variation in patient groups according to patient characteristics. The risk of postoperative infection after hip surgery may be higher in frail elderly than otherwise healthy young people. It can be relevant to identify subgroups of patients for whom specific cost-effectiveness analyses should be performed (30).

**Structural uncertainty** or model uncertainty goes beyond the uncertainty related to the parameters themselves and assesses the assumptions imposed by the modelling framework. The structural assumptions of the model and how these influence the estimated uncertainty should be considered (29).

### 1.2.5 Types of decision models for health economic evaluation

The basic tool used to solve clinical decision problems under uncertainty is a decision tree. Decision trees work well for analyses with limited recursion and limited time horizons. In a decision tree, uncertain events are represented by events at chance nodes. However, more advanced models may be needed when the decision problem involves recurring events, extended time horizons and when the timing of events is important(24). The purpose of a model is to inform medical decisions and resource allocation. Different areas such as clinical, epidemiological and economic data can provide the evidence base that is structured by use of quantitative methods to assist decision makers in making informed decisions. A model-based analysis is valuable not only in the ability to provide a point estimate in a treatment decision but also in the systematic examination and reporting of uncertainty around the decision (30).

It may be useful to distinguish between two overarching types of models. The first is **Patient-level stochastic simulations**, including discrete event simulations (DES) (32) and state-transition microsimulation (33). A key feature of these models is events occurring at the patient level, which require simulation of numerous individual patients. Assessment of parameter uncertainty in these

models require elimination of stochastic uncertainty (the term Monte Carlo error has been used in these cases) (30).

The second overarching category is **Markov cohort models** (33). In Markov models, a whole cohort of patients is followed in time simultaneously. In these models, the stochastic uncertainty does not have to be disentangled from parameter uncertainty (30). State-transition microsimulation has important features in common with Markov cohort models and both can be described as state-transition models (see state-transition models below).

Discrete event simulation provides a flexible framework including the ability to represent complex behaviour within, and interactions between individuals, populations and their environment. "Discrete" refers to the fact that a DES jumps forward in time at discrete intervals, from one event in time to the next. The term also refers to the discrete character of the events, which means that they are mutually exclusive. The DES typically represents an environment such as a hospital or a particular disease in a defined population, as for example patients with cardiovascular disease in Norway. The main building blocks in a DES are entities, attributes, events, resources, queues and time. In health care, the entities are typically patients but they might also be caregivers or items such as organs. The entities have attributes, experience events, consume resources and enter queues over time. The attributes are specific features that make it possible for the entity to carry information about age, sex, health status, quality of life and health care costs. Events are in broad terms the things that can happen to an entity or to the environment and can for example be a progression of the disease to a new stage, a hospital admission or a dose increase of a medication used. A resource is something that provides a service to an entity, for example a surgical room. If the resource is occupied when it is demanded of the entity, the entity enters a queue. Finally, time is an important component in DES. At start of the simulation, a simulation clock starts and keeps track of time spent in the model. The use of the clock makes it possible to count detailed periods such as time since diagnosis, hospital stay and symptom duration. All of these characteristics make the DES a very flexible tool. This flexibility allows for events occurring at any time and is particularly relevant when the time to each event is important in patients with multiple or competing risks. Many patient characteristics can be taking into account and they can change over time. When the disease process involves a series of events (e.g. myocardial infarction, resuscitation, percutaneous coronary intervention stenting and stroke) the DES is suitable. The DES was primarily constructed for solving scenarios when patients' demand for a particular resource and their priority status in a queue might be influenced by their attributes and is clearly a good choice for such problems. Further, the DES is recommended when the problem under study involves limited resources. Generally, the DES is used for representing complex systems and often requires extensive data. The programming can be made with general programming languages (e.g. C++ or R) but software specifically developed for DES is also available (32).

State-transition models are used for clinical situations which can be described in terms of the health conditions that individuals can be in ("states"), how the individuals move between the states ("transitions") and how probable such moves are ("transition probabilities"). In these situations state-transition microsimulation or Markov cohort models are suitable. The key features of these models are states, transitions, initial state vector, transition probabilities, cycle length, state values ("rewards"), logical tests performed at the beginning of each cycle to determine the criteria and termination criteria. The states are mutually exclusive which means that any individual can be in only one state during each cycle. Further, the states are collectively exhaustive which means that any

individual in the initial cohort must be in a state during each cycle. The states should reflect the disease/health process and should capture the benefits or harms of any interventions. The transitions among the states should reflect the expected disease progress with or without any intervention. State-transition microsimulation simulates only one individual at a time and the simulation is evaluated using first-order Monte Carlo simulation. This means that whether an individual facing a certain transition probability makes the transition depends on a random number. Cohort models are analysed as single cohorts making the progression through the states at the same time. At the start of a cohort simulation, a hypothetical cohort is allocated among the defined states (33). The length of the cycle can vary across models but should be chosen to reflect a clinically meaningful time interval, which implies that we can expect defined events preferably to occur only once during a cycle. Availability of data also influences cycle length (24). Shorter cycles will yield estimates that are more precise. In each state, values are assigned to yield the expected outcomes. Utilities can for example be assigned to derive QALYs. Similarly, costs corresponding to each state should be assigned. Termination criteria (*i.e*. number of cycles the model is supposed to run) are set to determine the time horizon of the analysis. The time horizon should be sufficiently long to capture all relevant outcomes in terms of health effects and costs related to the decision problem. A state-transition model is relatively simple to develop, debug and analyse given that the number of states is not extensive. The main disadvantage is the assumption that transition probabilities are independent of history. The transition probabilities are neither dependent on past states, nor time spent in the current state. This assumption is generally called the "Markovian" property (33).

A state-transition model is a recommended when the decision problem can be framed in terms of states, when interactions between individuals are not relevant and the population under study is a closed cohort. A state-transition model can be used for evaluating different types of interventions such as primary prevention, screening, diagnosis and treatment. In primary prevention, the decision problem concerns what happens before an individual gets a diagnosis/disease and the focus is on risk factors. In screening evaluations, state-transition models have evaluated both one-time screening and repeated screening programmes. Screening programmes can differ in several aspects as type and sequence of testing, screening interval, start, and stop dates for the programme. Diagnostic state-transition models have been used for identifying optimal strategies in individuals with symptoms or signs of a potential disease. The options may involve choice of different tests, one versus several tests, or the development of new diagnostic technologies. Treatment evaluation is restricted to anyone who already has a clinical condition or diagnosis. A suitable state-transition model disease process should mirror the disease's natural history, expected prognostic pathways without any intervention and effects of treatment (33).

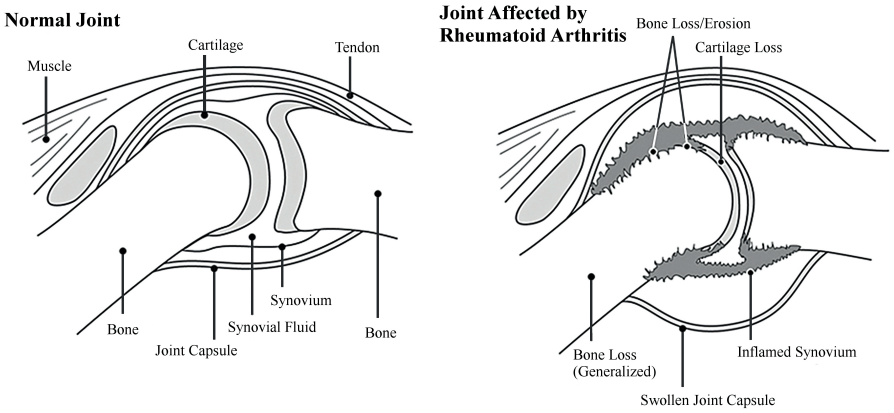
## 1.3 Rheumatoid arthritis

### 1.3.1 Symptoms and clinical findings

RA is a chronic inflammatory disease. The patients suffer from swelling, tenderness, pain and destruction of joints, causing disability and increased mortality (1-4). The disease is autoimmune. Pathogenic immune reactions, including antibody formation can be activated in genetically susceptible individuals. Autoantibodies typically include anti-citrullinated protein antibodies (ACPA) and rheumatoid factor (RF). These can be present in an individual many years before clinical symptoms of the disease are detected (1, 34).

RA entails both reversible and irreversible functional impairments. The reversible components of functional impairment include joint pain and swelling due to inflammation while the irreversible outcomes are related to joint destruction and deformity (35, 36). In figure 2, a normal joint and a joint affected by RA are presented. The affected joint exhibits inflammation and bone destruction.

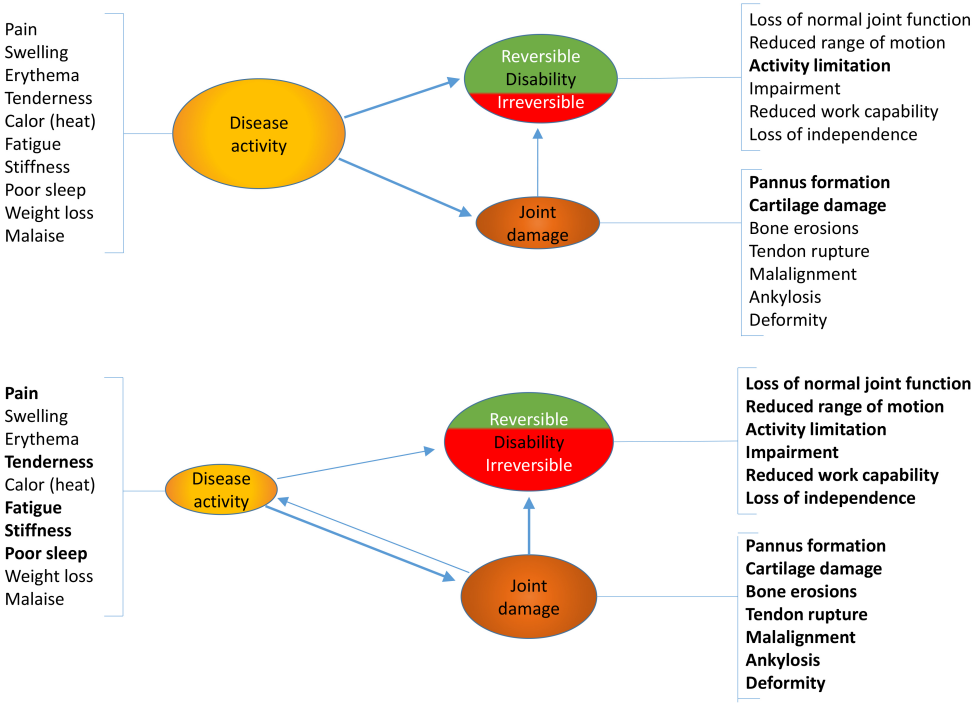
Figure 2. Normal joint and joint affected by RA.




National Institute of Arthritis and Musculoskeletal and Skin Diseases, US, 2013. Available from: http://www.niams.nih.gov/Health_Info/Rheumatic_Disease/default.asp

There is a clear link between inflammation, joint damage and physical function in RA. Thus, radiographic assessment of erosions and joint space narrowing is an important part of the assessment of patients. In early RA, inflammation and disease activity are responsible for most of the functional reduction (figure 3a). As the disease duration increases, joint damage is more closely related to disability (figure 3b) (37).

Figure 3a-b. Hypothesized link between disease activity, functional disability and structural joint damage in RA. Reprinted by permission from Macmillan Publishers Ltd: [Nature reviews Rheumatology] (Lillegraven S, van der Heijde D, Uhlig T, Kvien TK, Haavardsholm EA), copyright (2012) (37).




### 1.3.2 Classification

The American College of Rheumatology (ACR) published the old classification criteria in 1987-88 (table 4). These criteria were based on 262 patients who had RA for 7.7 years on average and a control group of 262 patients with other rheumatic diseases. The criteria included seven items, of which four had to be fulfilled for the classification of RA. Criteria one through four must have been present for at least six weeks (38). Classification criteria have been developed for research purposes and not for diagnosis. The classification criteria therefore represent a “classical case”. However, in clinical practise, classification criteria are also widely used as diagnostic criteria (39).

Table 4. The 1987 revised criteria for the classification of RA

| | Criterion | Definition |
|---|---|---|
| 1 | Morning stiffness | Morning stiffness in and around the joints, lasting at least 1 hour before maximal improvement |
| 2 | Arthritis of 3 or more joint areas | At least 3 joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints |
| 3 | Arthritis of hand joints | At least 1 area swollen (as defined above) in a wrist, MCP, or PIP joint |
| 4 | Symmetric arthritis | Simultaneous involvement of the same joint areas (as defined in 2) on both sides of the body (bilateral involvement of PIPS, MCPs, or MTPs is acceptable without absolute symmetry) |
| 5 | Rheumatoid nodules | Subcutaneous nodules, over bony prominences, or extensor surfaces, or in juxtaarticular regions, observed by a physician |
| 6 | Serum rheumatoid factor | Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in 4 % of normal control subjects |
| 7 | Radiographic changes | Radiographic changes typical of rheumatoid arthritis on posteroanterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify) |

Arthritis Rheum. 1988;31(3):315-24

The old classification criteria have been criticized for not being able to detect the disease in early stages. New classification criteria for RA were published in 2010, based on a joint working group from ACR and the European League Against Rheumatism (EULAR) (table 5). The new criteria were based on data from 3,115 patients with undifferentiated arthritis and disease duration of 1-7 months. The classification criteria are meant to test patients who have at least one joint with definite clinical synovitis, which cannot be better explained by another disease than RA. The criteria are based on an algorithm in which the scores from four categories (A-D) are added. A score of six or more out of ten classifies a patient as having RA (1, 39).

Table 5. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for RA

| | | Score |
|---|---|---|
| A | Joint involvement | |
| | 1 large joint | 0 |
| | 2-10 large joints | 1 |
| | 1–3 small joints (with or without involvement of large joints) | 2 |
| | 4–10 small joints (with or without involvement of large joints) | 3 |
| | >10 joints (at least one small joint) | 5 |
| B | Serology (at least 1 test result is needed for classification) | |
| | Negative RF *and* negative ACPA | 0 |
| | Low-positive RF *or* low-positive ACPA | 2 |
| | High-positive RF *or* high-positive ACPA | 3 |
| C | Acute-phase reactants (at least one test result is needed for classification | |
| | Normal CRP *and* normal ESR 0 | 0 |
| | Abnormal CRP *or* normal ESR 1 | 1 |
| D | Duration of symptoms | |
| | <6 weeks | 0 |
| | ≥6 weeks | 1 |

Ann Rheum Dis. 2010;69(9):1580-8

### 1.3.3 Epidemiology

The incidence was estimated at 25.7/100,000 (0.257‰) per year in Oslo during the period 1988 to 1993(40) and at 28.7/100,000 (0.287‰) in the county of Troms from 1987 to 1996 (41). The median incidence estimates varied from 0.1‰ to 0.5‰ in a systematic review based on the 1987 American College of Rheumatology Criteria for RA (42). Both prevalence and incidence rates of RA were reported lower in Southern European than in Northern European and North American countries (US), even though the differences were not statistically significant for females (42). The incidence of RA has been reported to diminish in the last decades (43).

The overall prevalence of RA in Norway was estimated at 0.44% for inhabitants between 20 and 79 years in Oslo (44) and at 0.39% in the county of Troms in Northern Norway (41). The median prevalence for North European countries was 0.50 % in the systematic review (42). A study by Simons *et al* report the prevalence rates in the United States to be 0.43% in 2006. Similar numbers are reported for 2004 (0.40%) and 2005 (0.44%).

RA is more common among women than men. Results from a study based both on a county register and a population survey in Oslo indicated a prevalence rate in females of 0.67% and 0.19% in males (44).

### 1.3.4 Treatment

The first International consensual treatment recommendations for RA were published in 2010 by the European League Against Rheumatism (EULAR). A task force and five subgroups conducted systematic literature reviews on five topics with the final aim of publishing evidence-based recommendations for drug management of patients with RA. The task force ended on three overarching principles and 15 recommendations for the management of RA with synthetic and biologic DMARDs (45, 46).The EULAR recommendations were updated in 2013 and the number of

recommendations was reduced from 15 to 14 (14). The updated version is presented below (table 6). The three overarching principles consider both treatment effect and costs (table 6) (14, 45).

Table 6. 2013 Update of the EULAR recommendations for the management of RA with non-biological and biological DMARDs.

| **Overarching principles** |
|---|
| A. Treatment of RA patients should aim at the best care and must be based on a shared decision between the patient and the rheumatologist |
| B. Rheumatologists are the specialists who should primarily care for RA patients |
| C. RA incurs high individual, societal and medical costs, all of which should be considered in its management by the treating rheumatologist |

| **Recommendations** |
|---|
| 1. Therapy with DMARDs should be started as soon as the diagnosis of RA is made |
| 2. Treatment should be aimed at reaching a target of remission or low disease activity in every patient |
| 3. Monitoring should be frequent in active disease (every 1-3 months); if there is no improvement by at most 3 months after the start of treatment or the target has not been reached by 6 months, therapy should be adjusted |
| 4. MTX should be part of the first treatment strategy in patients with active RA |
| 5. In cases of MTX contraindications (or early intolerance), sulfasalazine or leflunomide should be considered as part of the (first) treatment strategy |
| 6. In DMARD naïve patients, irrespective of the addition of glucocorticoids, csDMARDs monotherapy or combination therapy of csDMARDs should be used |
| 7. Low-dose glucocorticoids should be considered as part of the initial treatment strategy (in combination with one or more csDMARDs) for up to 6 months, but should be tapered as rapidly as clinically feasible |
| 8. If the treatment target is not achieved with the first DMARD strategy, in the absence of poor prognostic factors, change to another csDMARD strategy should be considered; when poor prognostic factors are present, addition of a bDMARD should be considered |
| 9. In patients responding insufficiently to MTX and/or other csDMARD strategies, with or without glucocorticoids, bDMARDs (TNF inhibitors*, abatacept or tocilizumab, and, under certain circumstances, rituximab†) should be commenced with MTX |
| 10. If a first bDMARD has failed, patients should be treated with another bDMARD; if a first TNF inhibitor therapy has failed, patients may receive another TNF inhibitor* or a biological agent with another mode of action |
| 11. Tofacitinib may be considered after biological treatment has failed |
| 12. If a patient is in persistent remission after having tapered glucocorticoids, one can consider tapering‡ bDMARDs§, especially if this treatment is combined with a csDMARD |
| 13. In cases of sustained long-term remission, cautious reduction of the csDMARD dose could be considered, as a shared decision between patient and physician |
| 14. When therapy needs to be adjusted, factors apart from disease activity, such as progression of structural damage, comorbidities and safety issues, should be taken into account |

Ann Rheum Dis. 2014;73(3):492-509.

*TNF inhibitors: adalimumab, certolizumab pegol, etanercept, golimumab, infliximab, biosimilars (as approved according to a thorough approval process, such as by EMA and/or FDA).
†The 'certain circumstances', which include history of lymphoma or a demyelinating disease, are detailed in the accompanying text.

‡Tapering is seen as either dose reduction or prolongation of intervals between applications.
§Most data are available for TNF inhibitors, but it is assumed that dose reduction or interval expansion is also pertinent to biological agents with another mode of action.
DMARD, disease-modifying antirheumatic drug; EMA, European Medical Agency; EULAR, European League against Rheumatism; FDA, Food and Drug Administration; MTX, methotrexate; RA, rheumatoid arthritis; TNF, tumour necrosis factor; bDMARD, biologic DMARD; csDMARD, conventional synthetic DMARD.

Treatment for RA comprises pharmacological treatment with DMARDs, non-steroidal anti-inflammatory drugs (NSAIDs) and glucocorticosteroids (GCs) as well as non-pharmacological interventions including physical-, and occupational therapy and psychological interventions. The cornerstone is the use of DMARDs (45, 47). DMARD is an overarching term of different antirheumatic agents which can affect the immune response or suppress the disease process (48), in contrast to symptom-modifying agents such as NSAIDs and simple analgesics. The term refers to historical treatments as for example injectable gold as well as more recently used drugs such as methotrexate, sulfasalazine or leflunomide (49). However, a new class of DMARDs, the so-called biologic DMARDs (bDMARDs) have become available during the last decade. Biological products include a broad range of products such as a vaccine, blood, allergenic, gene therapy, tissue, recombinant therapeutic protein and somatic cells (50). Biological processes rather than chemical synthesis produce biologics. In most cases, the term biologics is used for a class of therapeutics produced by biological processes involving recombinant DNA technology. Products produced by recombinant DNA include medications with different mechanism of action. The bDMARDs have different mechanisms of action, targeting the tumour necrosis factor, the interleukin 1 receptor, the interleukin 6 receptor, or B lymphocytes and T-cells (51). The effect of bDMARDs has been documented in randomized controlled trials (RCTs) as well as in systematic reviews (7-11).

### 1.3.5 Cost-of-illness

In the cost expressions in this thesis, except from the ones in the articles, exchange rates from xe.com 05.09.2013 have been used. During the last decade, the economic burden of RA has been assessed in several cost-of-illness studies. The methods used have varied, both in terms of patient characteristics, severity of disease and methodology for assessment of production losses (FCA or HCA) (52). In a systematic review, comprising 26 studies of the costs related to RA, the mean direct cost per patient per year was estimated at €4,170[IQR 2,756-4,561], the mean indirect cost at €1,441[IQR €702-€1,307] with the FCA and the mean indirect cost at €8, 452 [€4,144-€11,566] with the HCA (all costs in 2006 €). The overall mean total cost per year, assessed with the HCA was €14,906 and the share of productivity loss was estimated to 57% of the total costs (HCA).The review encompassed studies from 1988 to 2007 and 22 of the 26 included studies reported costs of TNF inhibitors (TNFi) (53).

In an assessment of costs for RA-patients from the US, health care expenditures declined from €9,268 ($12,224) in 2004 to €7,140 ($9,416.86) in 2006. The estimates were based on data from the Medical Expenditure Panel Survey (MEPS), including approximately 34,000 individuals per year (54). Another recent study, performed in South Korea was based on 196 patients and performed in 2009 (55). The estimated direct health related costs for RA-patients were €2,152 (KRW3,109,000) and the costs of sick leave and work loss were €2,309 (KRW3,337,000), yielding a total cost estimate of €4,461 (KRW6,446,000). The costs of productivity loss accounted for 52% of the total cost (55). The health care cost estimate from South Korea was approximately half the size of the one from the review (53) and one fourth of the estimates from the US study (54).

In another recent US study, covering the period from January 2005 to March 2009, the yearly treatment cost of TNFi for patients with RA, psoriasis, PsA and AS was assessed. The study was based on the IMF LifeLink™Health Plan Claims Database and the total number of included patients was 27,704, of whom RA-patients constituted 18,094. The costs included wholesale acquisition costs for a TNFi and administration costs for the same drugs. In RA these costs were estimated at ($14,314/$17,700/$20,390) (€10,853/€13,420/€15,460) depending on medication used (etanercept/adalimumab/infliximab) (56). A similar study from the US, also including patients between 2005 and 2009, but based on the Thomson Reuters MarketScan® Commercial Claims and Encounters Database, reported medication and administration cost estimates very close to the ones from the IMF LifeLink™Health Plan Claims Database. The study included 21,652 patients in total of whom 13,850 had RA. The estimates were: ($14,892/$18,381/$23,265) (€11,291/13,937/€17,640) depending on TNFi (etanercept/adalimumab/infliximab) (57).

Eriksson et al. estimated the mean annual health care costs for prevalent RA-patients in Sweden at €6,352 (€6,239 (SD 8,755) for patients 18-64 years and €6,438 (SD 9,392) for patients ≥65 years) in a register-based study published in 2013. The data sources were the National Patient Register, the Swedish Rheumatology Quality Register, the Prescribed Drug Register, the Social Insurance Agency and Causes of Death Register. The resource use was from the year 2010 and included non-primary outpatient care visits, hospital admissions and drug use. The mean annual cost was estimated at €23,147 (SD 23,099) with the HCA and €16,712 (SD 15,378) with the FCA when also productivity losses were included. The costs for an incident cohort of RA-patients were also estimated. The mean health care costs for the incident cohort was €4,623 (SD 6,370) for patients 18-64 years and €7,784 (SD 9,831) for patients ≥65 years). The costs of production loss were lower in the incident cohort than the prevalent cohort (58).

## 1.4 Psoriatic arthritis

PsA belongs to a group of inflammatory rheumatic joint diseases, named spondyloarthritis (SpA). Besides PsA, the group includes arthritis associated with inflammatory bowel disease, reactive arthritis, AS and undifferentiated SpA (59). The diseases in the group have overlapping features such as arthritis in the axial skeleton including inflammatory back pain, uveitis, gastroenterological and dermatological symptoms and a genetic link through HLA-B27 (60). SpA is often categorized into predominantly peripheral SpA or predominantly axial SpA, according to the clinical manifestations (59).

### 1.4.1 Symptoms and clinical findings

PsA is a chronic inflammatory disease affecting joints and skin (5). The disease is heterogeneous and peripheral arthritis, axial disease, skin and nail disease, dactylitis and enthesitis are typical features (61). Dactylitis means inflammation in the whole finger or toe and enthesitis refers to inflammation at the insertion of a ligament or tendon into the bone.

The disease often has a fluctuating course and inflamed joints are swollen, painful and tender. Rather frequent symptoms are joint pain as well as stiffness in the late night or morning with duration of over 30 minutes. If axial involvement is present, back pain in the night can occur. Skin manifestations often present years before the joints become involved. About 1/3 to 1/4 of psoriasis patients develop PsA (5).

### 1.4.2 Classification

In 2006, the CASPAR (ClASsification criteria for Psoriatic Arthritis) criteria were published (table 7). They were based on a prospective international study with the aims of 1) comparing the performance of previously existing criteria and 2) to examine if more accurate criteria could be derived from direct data examination. The study included 588 PsA cases and 536 controls (RA n=384, AS n=72, undifferentiated arthritis n=38, connective tissue disorders n=14 and other diseases n=28). The resulting CASPAR criteria had a specificity of 98.7% and a sensitivity of 91.4% (62).

Table 7. The CASPAR criteria

| To meet the CASPAR (ClASsification criteria for Psoriatic ARthritis) criteria, a patient must have inflammatory articular disease (joint, spine, or entheseal) with ≥3 points from the following 5 categories: |
|---|
| 1. Evidence of current psoriasis, a personal history of psoriasis, or a family history of psoriasis.<br>Current psoriasis is defined as psoriatic skin or scalp disease present today as judged by a rheumatologist or dermatologist.*<br>A personal history of psoriasis is defined as a history of psoriasis that may be obtained from a patient, family physician, dermatologist, rheumatologist, or other qualified health care provider.<br>A family history of psoriasis is defined as a history of psoriasis in a first- or second-degree relative according to patient report. |
| 2. Typical psoriatic nail dystrophy including onycholysis, pitting, and hyperkeratosis observed on current physical examination. |
| 3. A negative test result for the presence of rheumatoid factor by any method except latex but preferably by enzyme-linked immunosorbent assay or nephelometry, according to the local laboratory reference range. |
| 4. Either current dactylitis, defined as swelling of an entire digit, or a history of dactylitis recorded by a rheumatologist. |
| 5. Radiographic evidence of juxtaarticular new bone formation, appearing as ill-defined ossification near joint margins (but excluding osteophyte formation) on plain radiographs of the hand or foot. |

*Current psoriasis is assigned a score of 2; all other features are assigned a score of 1
Arthritis Rheum. 2006;54(8):2665-73

### 1.4.3 Epidemiology

Both prevalence and incidence of PsA seem to vary between different geographic areas. The incidence vary with rates from 0.1/100,000 (0.001‰) in Japan to 23/100,000 (0.23‰) in Finland (63). In a recent Norwegian study with data from the Nord-Trøndelag Health Study 3 (HUNT 3), the incidence rate was 41/100,000 (0.41‰) which is higher than previously observed rates (64).

A systematic review reported a variation in the prevalence estimates of PsA from 1/100,000 (<0.01%) in Japan to 420/100,000 (0.42%) in Italy.(63) In the Norwegian study from Nord-Trøndelag, the prevalence was found to be as high as 670/100,000 (0.67%)(64).

Different methodology may contribute to the variation in the reported frequency estimates. The lack of commonly accepted diagnostic criteria may be important. Some studies have used the European Spondylarthropathy Study Group (ESSG) criteria but others have not (63). The Norwegian study used the more recent CASPAR criteria (62). The prevalence studies have used both retrospective and cross-sectional approaches and the incidence studies have used both retrospective and prospective

designs (63). The reasons for the high frequency in central Norway is unknown and environmental and genetic factors, awareness of the disease as well as diagnostic criteria may have influenced the results (64).

PsA seems to occur with about the same frequency in males and females and average age at onset is 36–40 years in most studies (65).

### 1.4.4 Treatment

EULAR initiated a task force with the aim of developing recommendations for pharmacological treatment based on evidence from a systematic literature review and expert opinion, in the same way as had been done for RA. In PsA, 5 overarching principles and 10 recommendations were agreed upon (table 8) (66).

In the majority of PsA patients, the first-line treatment is NSAIDs, although the data are limited on the benefits of NSAIDs for PsA. NSAIDs have shown efficacy for joint symptoms, but not for the skin involvement. The second option is sDMARD therapy, which should be started in patients with active disease. However, the data on who should start DMARD therapy and when, are sparse and future research is recommended to address these issues. In PsA, as in RA, the first choice of DMARD should be MTX. Local injections of corticosteroids can be useful as adjunctive therapy (66, 67). If a patient has not reached a low disease-activity status after usually 3-6 months, TNFi treatment can be considered in patients with active disease. The definition of active disease in PsA should be addressed in future research (66). The efficacy of TNFi treatment in PsA has been documented in several trials, both in terms of reducing joint and skin symptoms as well as reducing structural changes, as seen by radiographic examinations (67-71).

Table 8. EULAR recommendations for the management of PsA, with levels of evidence, grade of recommendations and level of agreement

| **Overarching principles** | **Level of agreement (mean±SD)** |
|---|---|
| 1. Psoriatic arthritis is a heterogeneous and potentially severe disease, which may require multidisciplinary treatment. | 9.8±0.5 |
| 2. Treatment of psoriatic arthritis patients should aim at the best care and must be based on a shared decision between the patient and the rheumatologist. | 9.8±0.8 |
| 3. Rheumatologists are the specialists who should primarily care for the musculoskeletal manifestations of patients with psoriatic arthritis; in the presence of clinically significant skin involvement a rheumatologist and a dermatologist should collaborate in diagnosis and management. | 9.6±0.8 |
| 4. The primary goal of treating patients with psoriatic arthritis is to maximise long-term health-related quality of life, through control of symptoms, prevention of structural damage, normalisation of function and social participation; abrogation of inflammation, targeted at remission, is an important component to achieve these goals. | 9.7±0.6 |
| 5. Patients should be regularly monitored and treatment should be adjusted appropriately. | 9.7±0.7 |

| Recommendations | Level of evidence | Grade+ | Level of agreement (mean±SD) |
|---|---|---|---|
| 1. In patients with psoriatic arthritis, non-steroidal anti-inflammatory drugs may be used to relieve musculoskeletal signs and symptoms. | 1b | A | 9.4±0.9 |
| 2. In patients with active disease (particularly those with many swollen joints, structural damage in the presence of inflammation, high ESR/CRP and/or clinically relevant extraarticular manifestations), treatment with disease-modifying drugs, such as methotrexate, sulfasalazine, leflunomide, should be considered at an early stage. | *1b, †4 | B | 9.4±0.7 |
| 3. In patients with active psoriatic arthritis and clinically relevant psoriasis, a disease-modifying drug that also improves psoriasis, such as methotrexate, should be preferred. | 1b | A | 9.1±1.0 |
| 4. Local injections of corticosteroids should be considered as adjunctive therapy in psoriatic arthritis; systemic steroids at the lowest effective dose may be used with caution. | ‡3b, §4 | C | 8.9±1.2 |
| 5. In patients with active arthritis and an inadequate response to at least one synthetic disease-modifying antirheumatic drug, such as methotrexate, therapy with a tumour necrosis factor inhibitor should be commenced. | 1b | B | 8.9±1.5 |
| 6. In patients with active enthesitis and/or dactylitis and insufficient response to non-steroidal anti-inflammatory drugs or local steroid injections, tumour necrosis factor inhibitors may be considered. | 1b | B | 8.5±1.5 |
| 7. In patients with predominantly axial disease that is active and has insufficient response to non-steroidal anti-inflammatory drugs, tumour necrosis factor inhibitors should be considered. | 2b | C | 9.3±0.9 |
| 8. Tumour necrosis factor inhibitor therapy might exceptionally be considered for a very active patient naive of disease-modifying treatment (particularly those with many swollen joints, structural damage in the presence of inflammation, and/or clinically relevant extra-articular manifestations, especially extensive skin involvement). | 4 | D | 8.6±1.7 |
| 9. In patients who fail to respond adequately to one tumour necrosis factor inhibitor, switching to another tumour necrosis factor inhibitor agent should be considered. | 2b | B | 8.9±1.8 |
| 10. When adjusting therapy, factors apart from disease activity, such as comorbidities and safety issues, should be taken into account. | 4 | D | 9.5±1.0 |

Ann Rheum Dis. 2012;71(1):4-12

Recommendations with different levels of evidence within the recommendation are listed below.

[+]Categorization of levels of evidence and grades of recommendation

| | | |
|---|---|---|
| 1a | SLR of randomised controlled trials | A: directly derived from level 1 evidence |
| 1b | At least one randomised controlled trial | |
| 2a | At least one controlled study without randomisation | B: derived from 2 or extrapolated from 1 |
| 2b | At least one quasi-experimental study | |
| 3 | Descriptive studies (comparative, correlation, case-control) | C: derived from 3 or extrapolated from 1 or 2 |
| 4 | Expert opinion | D: derived from 4 or extrapolated from 1, 2 or 3 |

The level of agreement was computed as a 0 to 10 scale, based on 28 voters within the group.
* In patients with active disease (particularly those with many swollen joints—usually ≥5, structural damage in the presence of inflammation, high ESR/CRP and/or clinically relevant extra-articular manifestations), treatment with disease-modifying drugs, such as methotrexate, sulfasalazine, leflunomide, should be considered;
†at an early stage.
‡ Local injections of corticosteroids should be considered as adjunctive therapy in psoriatic arthritis;
§ systemic steroids at the lowest effective dose may be used with caution.
CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; EULAR, European League Against Rheumatism; PsA, psoriatic arthritis.

Since the publications of these recommendations a new bDMARD, ustekinumab, targeting IL-12 and IL-23, has been approved for use in PsA (72, 73). Five TNFi are now approved for treatment of PsA (adalimumab, certolizumab pegol, etanercept, golimumab, infliximab).

### 1.4.5 Cost-of-illness

The medical expenditures associated with PsA are considerable, especially for patients on bDMARD therapy (56, 57, 74). A recent study from the US examined health care costs for one year in patients with PsA (n=3,164) who started on a bDMARD in the period January 2005-December 2009. The patients started treatment with a bDMARD, either as monotherapy or in combination with sDMARD(s). The health care costs assessed were inpatient, outpatient, and emergency department in addition to pharmacy expenditures. There was no significant difference in the total costs between the bDMARD monotherapy and the bDMARD +sDMARD groups. For the two groups together, the 1-year health care costs amounted to $26,535 (€20,119)(74).

In another recent US study which included patients from January 2005 to March 2009, the yearly treatment cost of TNFi for patients with RA, psoriasis, PsA or AS was assessed. 3,738 PsA-patients were included in the study. In PsA these costs were estimated at ($15,030/$18,483/$24,974) (€11,396/€14,014/€18,936) according to medication used (etanercept/adalimumab/infliximab) (56). A similar study from the US, also including patients between 2005-2009, but based on the Thomson Reuters MarketScan® Commercial Claims and Encounters Database, reported medication and administration cost estimates very close to the ones from the IMF LifeLink™Health Plan Claims Database. The estimates were: ($15,790/$18,031/$26,973) (€11,972/€13,671/€20,451) with the TNFi (etanercept/adalimumab/infliximab) (57).

## 1.5 Ankylosing spondylitis

AS belongs to the SpA group. See 1.4 Psoriatic arthritis (59).

### 1.5.1 Symptoms and clinical findings

AS is an inflammatory rheumatic joint disease which affects the axial skeleton.(6) Inflammation in the sacroiliac joints is the hallmark of the disease. Syndesmophytes (formation of new bone) at the vertebral bodies which can form bridges (ankylosis) are also characteristic features of the disease (75) but are late manifestations which can be detected on conventional radiographs, usually some

years after the onset of the disease. Other features of the disease are low bone density, osteoporosis and an increased risk of fractures. Peripheral arthritis predominantly affects the lower limbs but also hip and shoulder joints may be inflamed.(6) Spinal stiffness and reduced spinal mobility are characteristic symptoms. Inflammation, structural damage or both cause these symptoms (76). In addition, the disease may affect the eyes (uveitis) which indicates that, AS is a generalized disease of the body.

### 1.5.2 Classification

The classical New York AS criteria from 1966 were last modified in 1984 (table 9) (77). They have been used not only for classification but also for diagnosis of the disease (6).

Table 9. Modified New York criteria for ankylosing spondylitis

| |
|---|
| A. Diagnosis |
| 1. Clinical criteria<br>a) Low back pain and stiffness for more than 3 months which improves with exercise, but is not relieved by rest.<br>b) Limitation of motion of the lumbar spine in both the sagittal and frontal planes.<br>c) Limitation of chest expansion relative to normal values corrected for age and sex. |
| 2. Radiologic criterion<br>Sacroiliitis grade ≥2 bilaterally or Sacroiliitis grade 3-4 unilaterally. |
| B. Grading |
| 1. Definite ankylosing spondylitis if the radiologic criterion is associated with at least 1 clinical criterion. |
| 2. Probable Ankylosing spondylitis if:<br>a) Three clinical criteria are present.<br>b) The radiologic criterion is present without any signs or symptoms satisfying the clinical criteria. (Other causes of sacroiliitis should be considered.) |

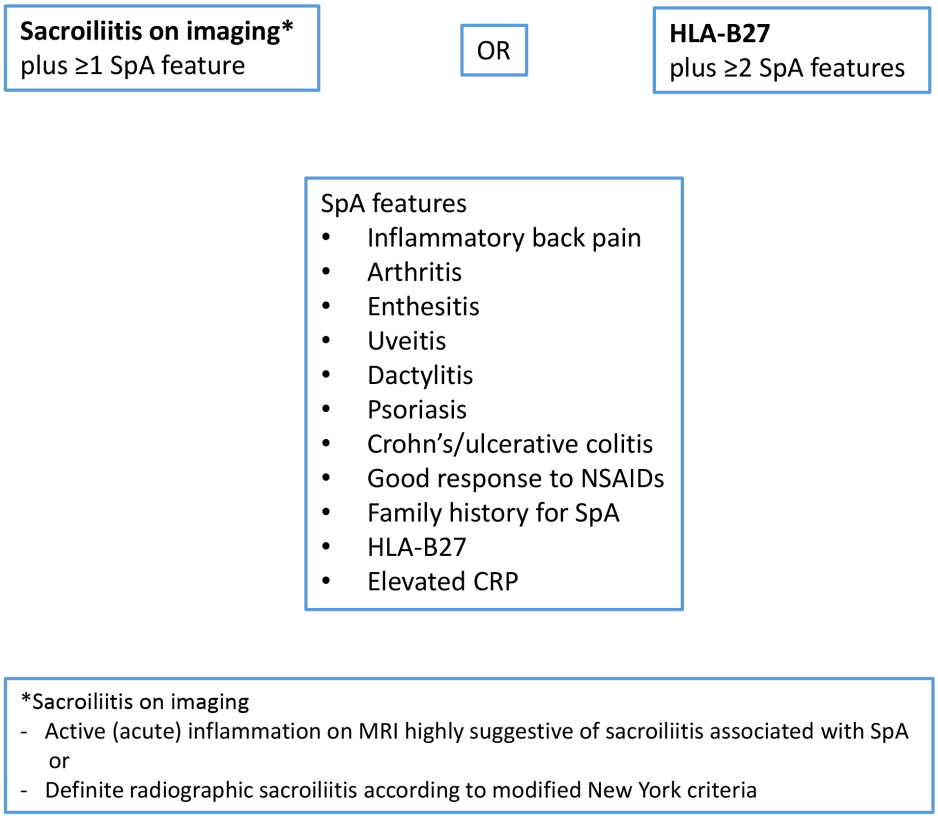
Arthritis Rheum. 1984;27(4):361-8

With the aim of being able to detect the disease earlier, in particular in patients with mainly axial or peripheral symptoms, two more clinically based sets of criteria were developed. One was the European Spondyloarthropathy Study Group (ESSG) criteria (78) and the other was the Amor criteria (79). However, these are both classification criteria, and work well as such but have additionally been used as diagnostic criteria. Use of classification criteria for diagnostic purposes can lead to an over- or underestimation of the frequency of the disease (6).

In 2004, the Assessment of Spondyloarthritis international Society (ASAS) decided to improve the Spondyloarthritis criteria especially for use in early disease. In the first step, the ASAS group focused on patients with predominantly axial SpA. A new set of classification criteria for axial SpA was published in 2009. Two sets of predefined criteria were compared to an expert physician diagnosis and to the ESSG and Amor criteria. The new criteria were first tested in 649 new patients from 25

centres included in the ASAS. The patients had chronic (≥3 months) back pain of unknown origin that started before 45 years of age. The reference standard was the clinical diagnosis made by the ASAS rheumatologist (SpA or no SPA). Refinement of the candidate criteria resulted in two new complementing criteria; the "imaging arm" and the "clinical arm" (figure 4). The sensitivity and specificity of the new double set of criteria were 82.9% and 84.4% respectively. The specificity of the new criteria was better than that of the ESSG and Amor criteria. The new criteria provide a standard for classifying non-radiographic axial SpA, which is important for reliable classification of patients in clinical trials and observational studies (59).

Figure 4. ASAS classification criteria for axial SpA. Patients with back pain (≥3 months) and age at onset <45 years (adapted from Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. Ann Rheum Dis(59)).




### 1.5.3 Epidemiology

An Incidence rate of 10.6/100,000 (0.106‰) has been reported from Northern Norway (80) and a rate as low as 0.5/100,000 (0.005‰) in Japan (81).

The prevalence of AS has been reported at 1.1-1.4% and 1.8% in population studies in Northern Norway but a newer study from the same region reported a prevalence of 0.4%, based on hospital records (60, 80). In a recent Swedish study, the prevalence of AS was estimated to be 0.12% (82) The

Swedish prevalence estimate was very close to estimates reported from Iceland and from Finland (0.13% and 0.15%, respectively) (60).

The male to female prevalence ratio is approximately 2:1(6). In earlier studies, men were more often affected than women were but some recent studies indicate a 1:1 ratio. There has probably been an under diagnosis of female AS patients in earlier decades because women have a slower disease progression and because AS is less common in females (83). The disease generally presents around 25 years of age (6) and rarely after 45 years of age (84). A positive HLA-B27 test is associated with an increased risk of SpA (85) and approximately 90% of AS patients are positive for HLA-B27 (84).

### 1.5.4 Treatment

ASAS developed management recommendations for AS in collaboration EULAR in 2006 (86), with an update in 2011 (87). The methodology was the same as for other rheumatic diseases. A systematic literature review was performed and constituted the basis for expert discussions and a consensus process. The recommendations are meant for patients who fulfil the New York criteria for AS. The updated version includes 4 overarching principles and 11 recommendations (table 10) (87).

Table 10. 2011 update of the ASAS/EULAR recommendations for the management of AS

| | |
|---|---|
| The overarching principles of the management of patients with AS are: | |
| • AS is a potentially severe disease with diverse manifestations, usually requiring multidisciplinary treatment coordinated by the rheumatologist. | |
| • The primary goal of treating the patient with AS is to maximise long term health-related quality of life through control of symptoms and inflammation, prevention of progressive structural damage, preservation/normalisation of function and social participation. | |
| • Treatment of AS should aim at the best care and must be based on a shared decision between the patient and the rheumatologist. | |
| • The optimal management of patients with AS requires a combination of non-pharmacological and pharmacological treatment modalities. | |
| **1.** | **General treatment**<br>The treatment of patients with AS should be tailored according to: |
| | • The current manifestations of the disease (axial, peripheral, extra-articular symptoms and signs).<br>• The level of current symptoms, clinical findings and prognostic indicators.<br>• The general clinical status (age, gender, comorbidity, concomitant medications, psychosocial factors). |
| **2.** | **Disease monitoring**<br>The disease monitoring of patients with AS should include: |
| | • Patient history (e.g., questionnaires)<br>• Clinical parameters<br>• Laboratory tests<br>• Imaging<br>• All according to the clinical presentation as well as the ASAS core set<br><br>The frequency of monitoring should be decided on an individual basis depending on:<br>• Course of symptoms<br>• Severity<br>• Treatment |

| | |
|---|---|
| **3.** | **Non-pharmacological treatment** |
| | • The cornerstone of non-pharmacological treatment of patients with AS is patient education and regular exercise.<br>• Home exercises are effective. Physical therapy with supervised exercises, land or water based, individually or in a group, should be preferred as these are more effective than home exercises.<br>• Patient associations and self-help groups may be useful. |
| **4.** | **Extra-articular manifestations and comorbidities** |
| | • The frequently observed extra-articular manifestations, for example, psoriasis, uveitis and IBD, should be managed in collaboration with the respective specialists.<br>• Rheumatologists should be aware of the increased risk of cardiovascular disease and osteoporosis. |
| **5.** | **Non-steroidal anti-inflammatory drugs** |
| | • NSAID, including Coxibs, are recommended as first-line drug treatment for AS patients with pain and stiffness.<br>• Continuous treatment with NSAID is preferred for patients with persistently active, symptomatic disease.<br>• Cardiovascular, gastrointestinal and renal risks should be taken into account when prescribing NSAID. |
| **6.** | **Analgesics** |
| | • Analgesics, such as paracetamol and opioid (like) drugs, might be considered for residual pain after previously recommended treatments have failed, are contraindicated, and/or poorly tolerated. |
| **7.** | **Extra-articular manifestations and comorbidities** |
| | • The frequently observed extra-articular manifestations, for example, psoriasis, uveitis and IBD, should be managed in collaboration with the respective specialists.<br>• Rheumatologists should be aware of the increased risk of cardiovascular disease and osteoporosis. |
| **8.** | **Disease-modifying antirheumatic drugs** |
| | • There is no evidence for the efficacy of DMARD, including sulfasalazine and methotrexate, for the treatment of axial disease.<br>• Sulfasalazine may be considered in patients with peripheral arthritis. |
| **9.** | **Anti-TNF therapy** |
| | • Anti-TNF therapy should be given to patients with persistently high disease activity despite conventional treatments according to the ASAS recommendations.<br>• There is no evidence to support the obligatory use of DMARD before or concomitant with anti-TNF therapy in patients with axial disease.<br>• There is no evidence to support a difference in efficacy of the various TNF inhibitors on the axial and articular/entheseal disease manifestations; but in the presence of IBD a difference in gastrointestinal efficacy needs to be taken into account.<br>• Switching to a second TNF blocker might be beneficial especially in patients with loss of response.<br>• There is no evidence to support the use of biological agents other than TNF inhibitors in AS. |

| | |
|---|---|
| **10.** | **Surgery** |
| | • Total hip arthroplasty should be considered in patients with refractory pain or disability and radiographic evidence of structural damage, independent of age.<br>• Spinal corrective osteotomy may be considered in patients with severe disabling deformity.<br>• In patients with AS and an acute vertebral fracture a spinal surgeon should be consulted. |
| **11.** | **Changes in the disease course** |
| | • If a significant change in the course of the disease occurs, other causes than inflammation, such as a spinal fracture, should be considered and appropriate evaluation, including imaging, should be performed. |

Ann Rheum Dis. 2011;70(6):896-904

### 1.5.5 Cost-of-illness

In AS, fewer cost-of-illness studies have been published than for RA (53). The review of cost-of-illness studies and economic evaluations by Franke and co-workers identified seven studies reporting on bottom-up costs in AS-patients. The review encompassed studies from 1998-2007 but none included costs of TNFi (53). The reported cost estimates are likely to be low compared to real costs because the treatment now frequently include TNFi in patients with high disease activity despite conventional therapy.(87) The mean direct cost estimate was €1,992[IQR €1,359-€2,474] and the mean annual productivity loss was €2,271[IQR €1,572-€2,970] with the FCA and €6,278 [€5,111-€7,725] with the HCA (all costs in 2006 €). In total, additionally including patient and family costs, the mean cost per year was estimated at €9,374 per patient when applying the HCA. The productivity costs constituted 66% of the total costs (53).

In the study based on 2012 claims data from the IMF LifeLink™Health Plan in the US, the number of included AS patients was 1,296. In AS the annual treatment cost of TNFi were estimated at $14,254/$16,925/$23,056 (€10,808/€12,833/€17,481) according to medication used (etanercept/adalimumab/infliximab) per patient (56). The corresponding numbers from the Thomson Reuters MarketScan® Commercial Claims and Encounters Database, were $14,348/$15,303/$24,945) (€10,879/€11,603/€18,914) with the same TNFi (etanercept/adalimumab/infliximab) (57). These numbers show that only the yearly medication cost for patients on TNFi-treatment are in the range €10,000-€19,000 in recently updated figures from the US Compared to the review from Europe, before the initiation of TNFi, the direct costs estimate per year was €1,992, also including other treatments than medications used (53).

## 1.6 Outcome measurements for inflammatory rheumatic joint diseases

Measurement in medicine can be done to fulfil three main objectives: to classify, to prognosticate and to assess change over time. Validity, reliability and responsiveness are important to consider when measuring change over time. In trials, the objective is to measure changes and differences, *e.g.* to evaluate the effect of treatment.(88) In a full economic evaluation, both consequences and costs of treatments are assessed. Thus, the effectiveness of the treatment is one of the included components (18). Both disease specific and generic measures of health are available. Outcome measures in inflammatory rheumatic joint diseases include measures of disease activity (89, 90), physical function (35, 91, 92), general concepts, health profiles, MAU instruments and scores for abnormalities for varying imaging modalities.

### 1.6.1 Disease specific measures in rheumatoid arthritis

The disease specific measurements in this section are limited to instruments that have been used in economic evaluations of health care interventions for RA.

One widely used measure of disease activity is the Disease Activity Score (DAS)(93, 94) and the further development of this score; the modified DAS that include 28-joint counts (DAS28) (95). The original DAS includes the Ritchie Articular Index and assessment of 44 joints with regard to swelling. Additionally it includes the erythrocyte sedimentation rate (ESR) and a general health assessment scored on a visual analogue scale (VAS). The more frequently used DAS28 reduced the number of included joints to 28, with separate assessments of tenderness and swelling. It was found to have comparable validity to disease activity scores with more joints included in the assessment (95). The DAS28 score ranges from 0 to 10 where a lower score indicates a lower disease activity (96). Remission has been defined at DAS28 <2.6 and cutpoints for low (≥2.6 and ≤3.2), moderate (>3.2 and ≤5.1) and high (> 5.1) disease activity have been identified (97, 98).

A patient reported measure of physical function is the Stanford Health Assessment Questionnaire (HAQ) (91). The HAQ has been extensively used, both in clinical practise and clinical trials in RA (99). The Modified Health Assessment Questionnaire (MHAQ) is a shortened version of the HAQ (92). The original HAQ includes eight categories of activities of daily living, with two to three sub questions in each domain, yielding in total 20 items. In MHAQ, there is only one question for each domain, reducing the total number of questions to eight. Patients respond to the questions by answering in one of four levels: “without any difficulty”= 0, “with some difficulty” = 1, “with much difficulty” = 2 or “unable to do” = 3, both in the HAQ and MHAQ (92). The two versions have been examined for differences, and the results indicated that there is a stronger ceiling effect in MHAQ than in HAQ, there were differences in assessments of patients with high levels of disability and finally, MHAQ in general provided lower scores (100).

### 1.6.2 General concepts

The concepts of Patient Acceptable Symptom State (PASS) and Minimal Clinically Important Improvement (MCII) are used for reporting the proportion of patients in an acceptable state and the proportion of patients who have experienced an important improvement in their condition (101). The concepts were discussed at Outcomes Measures in Rheumatology (OMERACT) 8 and a following survey confirmed the relevance of PASS and MCII in outcome assessment in rheumatic diseases (102). The EULAR and ACR have in a common task force underlined the importance of reporting improvement as well as sustainability of an acceptable level of the disease (89).

PASS and MCII are usually assessed by global anchoring questions about perceived state and change. Three methods have been used for assessing levels/changes in patient reported outcomes and clinical measures that correspond to PASS/MCII. These methods include the 75$^{th}$ percentile approach and two approaches using receiver-operating characteristic (ROC) analyses (103-105). One of the ROC approaches is maximum accuracy (103, 104) and the other 80% specificity (106). MCII is a rating of the extent to which patients’ health conditions have improved, and this type of rating can be called a transition rating. The validity of MCII has been suggested to be tested according to a method of transition ratings (107).

### 1.6.3 Generic measures of health

In this section, the presented instruments are limited to those included in the observational data available for the examinations in the presented thesis, *i.e.* in the Norwegian DMARD register study (NOR-DMARD)(108) and Oslo RA register (ORAR) (44). The included instruments were the Medical Outcomes Study 36-item Short-Form Health Survey (SF-36), the Short Form-6 Dimensions (SF-6D) and the EuroQol-5 Dimensions (EQ-5D) (109-111).

Generic instruments aim at being relevant for many patient groups and provide a means of making comparisons across patient groups. The generic measures of health have a standardized system for describing health or how health impacts on quality of life and some have an algorithm for valuing the described health states. The instruments, which include an algorithm, generate a score for each health state with full health represented by one and death by zero. These scores can subsequently be used to calculate QALYs (112). A commonly used name for this type of instruments is multi-attribute utility (MAU) instruments. MAU instruments commonly used in RA include the SF-6D (110) and the EQ-5D (111). The SF-6D is derived from the SF-36 (109).

The SF-36 is not a MAU instrument. It belongs to a class of instruments often called health profiles. It was developed to survey health status and includes 36 items divided in eight health concepts: physical functioning, social functioning, role limitations due to physical and mental health problems, bodily pain, mental health, vitality and general health (109). The SF-36 has been tested for reliability and validity in patients with various arthritides (113-116). In a study including 1016 patients with arthritis, the eight scales' reliability coefficients were in the rage of 0.75-0.91, exceeding the minimum reliability standard for group-level comparisons (0.70) estimated by Cronbach's alpha coefficients (113). The SF-36 scales were able to discriminate between different severity levels in arthritis at one point in time and to separate patients who responded from those who did not respond to treatment after two weeks. The instrument was valid for discriminating the effect of different treatment options in arthritis (115). In another study of 233 patients with RA, SF-36 was reliable (interclass coefficients 0.76-0.93) and the two summary mental and physical scales were valid and responsive for measuring health status in the patient group (114). In the Norwegian study, all scales' Cronbach's alpha coefficients exceeded 0.70 and the construct validity of the questionnaire was found to be satisfactory(116). The responsiveness to treatment has been further tested in SF-36 and the instrument was overall found to be responsive to treatment in RA (117, 118).

The SF-6D is a MAU-instrument, composed of six dimensions of health derived from the eights dimensions of the SF-36 (119). An algorithm computes a utility score from zero to one in SF-6D, and with the current algorithm no health state has a value below 0.296 except for death (110, 119). The SF-6D has been tested for reliability in patients with RA and the intraclass correlation coefficient (ICC) was 0.89 (95% CI:0.79,0.94) (120). SF-6D is also responsive to improvements in RA patients (120, 121).

EQ-5D is another MAU-instrument, developed by the EuroQol Group, a multidisciplinary research group with representatives from countries all over the world (122). The name reflects the composition of the group at the time of constitution. At present, EQ-5D is by far the most widely used generic HRQoL instrument worldwide (123). EQ-5D describes health along five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension can be rated at three levels, roughly corresponding to (1) no problems, (2) some problems, and (3) extreme problems. In total, 243 combination health states are possible (124). A new five-level version,

referred to as EQ-5D-5L, with new levels between the old ones, is currently being rolled out (125). The 5L-version is expected to replace the old 3L-version over time. In addition to the descriptive system, the instrument includes a global health assessment on a visual analogue scale (VAS), usually referred to as the EQ-VAS (124). Utility weights for EQ-5D health states have typically been assigned by asking members of the general public to value hypothetical health states using the time trade-off valuation method. National value sets, or tariffs, have been made in several countries. However, at the time of writing, no Norwegian tariff has been made. In countries without national tariffs, the UK tariff is the most commonly used. Health state values in the UK tariff range from -0.594 to 1.0 with 35% of all possible states below zero (126). The EQ-5D has rather low but acceptable reliability for group level comparisons, with ICC reported at 0.79 (95% CI:0.68,0.87) (118). In another study, the ICC was unacceptable low; 0.46 (95% CI:0.18,0.68) (120). Further, EQ-5D was found to have validity in being able to discriminate between low, moderate and high DAS28, and have acceptable responsiveness as reflected by analysis of mean score changes by one-way analysis of variance (ANOVA) and standardized response mean (SRM) estimates for change in health status (118).

## 1.7 Motivation for the development of the Norwegian Rheumatoid Arthritis (NORA) model

Current treatment recommendations suggest bDMARDs for RA-patients responding insufficiently to MTX and/or other sDMARD strategies. bDMARDs (TNF inhibitors, abatacept or tocilizumab, and, under certain circumstances, rituximab) are recommended to be commenced with MTX (table 6).

In Norway, the financing of drugs previously differed depending on whether the drugs were used in conjunction with hospitalisation or outpatient treatment and if the patient could administer the drugs at home. If the drugs were used in conjunction with hospitalisation or outpatient treatment, the hospitals were responsible for the costs of drugs used. If the patients could take the drugs at home, the National Insurance Scheme covered the costs. In 2006, the financing of all bDMARDs for inflammatory rheumatic joint diseases was transferred from the National Insurance Scheme to the Regional Health Authorities, which are responsible for the hospitals, for three main reasons:

- A possibility for inappropriate consumer shifts because different financing can lead to a choice of treatment based on economic and not medical considerations.
- Better priorities: The specialists (employed by the Regional Health Authorities) initiate bDMARD treatment and have the best qualifications for making treatment priorities.
- There was no price competition and it is allowed and possible to have a price competition between drugs with equivalent effectiveness and safety on a group level, for example by tender (127).

Since the costs of using bDMARDs are high and there was no price competition, a concern of the total costs of these drugs led to the implementation of a price bidding system in 2007. The pharmaceutical companies are now invited to participate in a yearly price bidding competition. The overall effectiveness and safety is assumed to be equal for TNFi and other bDMARDs on a group level. The result is a list of recommended therapies for patients with inflammatory rheumatic joint diseases, distributed to the four Regional Health Authorities in Norway.

Since no cost-effectiveness/cost-utility evaluation of treatment with TNFi+MTX versus sDMARDs for RA-patients had been performed previously in Norway, this was considered a relevant research area.

### 1.7.1 Economic evaluation of treatments for RA

Before we undertook an economic evaluation of bDMARDs for inflammatory rheumatic joint diseases in Norway, we wanted to assess previous research in this area. Therefore we started by performing a literature review.

#### *1.7.1.1 A literature review*

A literature search was performed with the aim of assessing previous economic evaluations of treatments for RA. Four groups of topics were identified for the search in Ovid MEDLINE 1948 to March Week 4 2011. Additionally, the same group of topics was identified in EMBASE 1980 to 2011 Week 13.

1. Disease: Arthritis, Rheumatoid (MeSH, focus) Medical Subject Headings (MeSH) is the US National Library of Medicine controlled vocabulary thesaurus used for indexing articles for PubMed.
2. Economic evaluation (Cost-Benefit Analysis MeSH or cost utility text word or cost effectiveness text word)
3. Model* (textword)
4. Antirheumatic Agents (MeSH) or Dmard* (textword)

Search date 05.04.2011 and updated search date 21.06.2012. Result: 58 at the first date and 66 at the second date, from which 28 were selected in MEDLINE. In EMBASE, the result was 53 on 05.04.2011. Finally, Cochrane Library was searched for the same groups of topics, yielding a result of 21 articles on 05.04.2011. After the second search, the search strategy was saved for MEDLINE and EMBASE and monthly updates of the search in MEDLINE and EMBASE were reported by email to the author.

#### *1.7.1.2 Results of the literature review*

In 1996, a US decision analytic model was developed for the evaluation of therapy with a potential biological agent for RA, compared to MTX and intramuscular gold (IMG) (128). This early model was based on a decision tree with three outcomes: no clinical response, partial response and complete response. The model was deterministic and one-way and two-way sensitivity analyses were performed. The study indicated that the cost of pharmaceuticals should be an important consideration for the development and use of biologic therapies for patients with RA (128). In 2002, a Markov model was used to estimate lifetime costs, life expectancy and quality-adjusted life expectancy for patients after use of infliximab plus MTX or MTX monotherapy for 54 weeks (129). The cost-effectiveness study took advantage of data from the randomized Anti-TNF Trial in Rheumatoid Arthritis with Concomitant Therapy (ATTRACT) (8). The results after 54 weeks were extrapolated in a Markov model including 21 health states, which combined five treatment options and four disability levels from the Health Assessment Questionnaire (HAQ) plus death in a lifetime perspective. The model had been constructed based on data from the Arthritis, Rheumatism and Aging Medical Information System (ARAMIS) (130). A societal perspective, including indirect costs was applied and the conclusion was that infliximab plus MTX for 54 weeks is likely to be cost-effective. A continuation of treatment beyond 54 weeks might also be cost-effective but it was concluded that longer follow-up clinical studies are needed to confirm the results (129).

A cost-effectiveness study performed in Europe examining the cost-effectiveness of infliximab plus MTX versus MTX alone (131) was also based on data from the ATTRACT trial (8). Unlike the previously

performed cost-effectiveness analysis using the same treatment data,(129) the European study used cohort studies from Sweden and the UK for the extrapolation of the clinical data in a 10-year Markov cohort model (131). The model was constructed with one-year cycles and six states based on the HAQ and one for death in each of the two treatment arms. EQ-5D was used for the estimation of QALYs and both direct and indirect costs were assessed. Disease progression was based on annual HAQ scores in the two cohort studies and transition probabilities between states were estimated using a regression model. The regression function controlled for age, gender and the time since disease onset and was used to generate transition probabilities for a cohort that matched the patients included in ATTRACT in terms of the predetermined characteristics. In Sweden the incremental cost per QALY of one and two years of infliximab therapy were €28,600 and €44,500 respectively while the corresponding cost per QALY were €41,500 and €56,100 in the UK when only direct costs were included. For both countries these ratios were reported to be in the usual range for treatments to be recommended (131).

Another Markov cohort model, developed in the Netherlands in 2004, estimated the five-year cost effectiveness of five different treatment strategies, including TNFi and leflunomide. The five options were examined with a specific Markov model for each treatment but the models were based on the same structure. The models had a cycle length of 3 months and the model ran for 5 years (20 cycles). Four Markov states were defined by the DAS; remission, low, moderate and high disease activity. The same cost and utility values of the states were used but specific treatment probabilities and costs of treatments were used in the different models. The authors concluded the strategy with the most favourable ICER among treatments including TNFi agents was as follows: Treatment with leflunomide, in the case of nonresponse after 3 months, switch to TNFi treatment, in the case of nonresponse to TNFi treatment after 3 months, switch to usual treatment. This strategy had a cost per QALY of €163,556 in comparison to usual treatment. Usual treatment in the Netherlands at the time of the study was to start with sulfasalazine and if there was insufficient effect or toxicity continue with MTX. If MTX was also ineffective or lead to toxicity several other classic DMARDs might be used (132).

The Birmingham Rheumatoid Arthritis Model (BRAM)(133) was developed in UK in 2004, building on an earlier preliminary version from 2002; the Birmingham Preliminary Model (BPM) (134). The focus was to evaluate two TNFis, etanercept and infliximab for the treatment of RA. Since the comparators in early trials of TNFis had been placebo and thus not reflected clinical practice this issue was addressed. Two separate analyses were compared, TNFi versus placebo and a sequence using TNFi to a sequence including current practice in the UK at the time. The authors used an individual sampling model. When placebo was used as comparator, the cost-effectiveness ratios were always lower compared to when an appropriate drug sequence was used. The advantage with the BRAM of modelling an appropriate comparator sequence of drugs is that the model can avoid new treatments' ICERs to appear lower than they really (133).

Another UK model for RA-patients, produced in Sheffield was published in 2004 (135). The objective was to assess the cost-effectiveness of etanercept monotherapy in patients previously failing two DMARDs, one of which was MTX, in line with the guidelines from the British Society for Rheumatology. The comparator was a sequence of traditional sDMARDs which was the current UK practice. The evidence base for the analysis was assembled from a range of sources; a review of trials and observational data. The authors stated that the main limitation of the study was availability of

data. The patient-level simulation model focused on the progression of HAQ score over time. To yield the QALY estimates, HAQ scores were converted to QALYs by means of a regression of HAQ to EQ-5D utility. Both direct and indirect costs were assessed but only direct costs were considered in the base-case. In a life-time perspective, the use of etanercept was found to be cost-effective as third-line therapy compared to a sequence of traditional DMARDs (135).

A new Markov cohort model for RA-patients was developed in Sweden in 2005, including both functional status as defined by the HAQ and disease activity as defined by the DAS28. 10 states where defined, first divided into five functional states which were further subdivided into high and low disease activity (136). The model was based on treatment effectiveness data over two years from the Trial of Etanercept and Methotrexate with Radiographic Patient Outcomes (TEMPO) (7). Cost and utility data were obtained from a survey of 616 patients with RA in Sweden. The model was run for 10 years and used one-year cycles. Transitions for all patients in the clinical trial were used directly during the first two years. Thereafter patients stopped treatment and followed the annual disease progression with standard treatment (+0.03 HAQ, based on a study by Scott et al(137)). In the base case with two years' treatment, the incremental cost per QALY for etanercept and methotrexate compared with methotrexate alone was estimated at €37,331. If a threshold value of €50,000 for the willingness-to-pay for a QALY was used, the probability that etanercept and methotrexate for two years was acceptable was 88%. In a 10-year perspective, assuming that patients continued treatment, the cost-effectiveness ratio per QALY gained was €46,494. The probability that the treatment was cost-effective over 10 years was 71% (136).

Another Markov cohort model was developed to estimate the cost-effectiveness of infliximab plus methotrexate versus methotrexate alone from a UK National Health Service (NHS) perspective. The model used clinical effectiveness data from the ATTRACT (8) for the first year of therapy, transition probabilities and health state valuations based on data from the ARAMIS (130) and resource use and costs from different sources in the UK. Only direct costs were included. Six-month cycles and 21 states were used in the model. The states were based on four disability categories from the HAQ, which were combined with five treatment options plus one death state. Four analyses were performed varying the treatment period with infliximab from one year to lifetime. In the primary analysis, considering 54 weeks treatment with infliximab plus methotrexate versus methotrexate alone, the incremental cost per QALY was £33,618 (€39,071). Assuming lifetime treatment with infliximab the incremental cost per QALY was reduced to £23,936(€28,459) due to a substantial QALY gain. The incremental cost per QALY always remained within the range for health care interventions normally funded in the UK (138).

Building on the previously described state-transition microsimulation model (135), a new model was developed with the objective of analysing the cost utility of adalimumab versus traditional DMARDs and other TNFi for RA-patients in Sweden. The analysis simulated 10,000 hypothetical patients with RA in 6 months cycles within a lifetime perspective. In each cycle, risk of withdrawal, adverse events and mortality were assessed. Treatment effectiveness data came from a review of the literature and two RCTs for adalimumab. The Health Utility Index-III (HUI-3) was used both as direct and indirect health utility measure. The trials with adalimumab included the HUI-3 but a transformation had to be performed from the HAQ to the HUI-3 for utility calculations for infliximab and etanercept. Swedish unit costs and treatment recommendations were used. The results indicated that the ICERs for the TNFi versus sDMARDs were between €35,000 and €65,000 and that adalimumab is at least as cost

effective as other TNFi. The results were in the range normally considered cost effective at that time in other European countries (139).

A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of RA and an economic evaluation of their effectiveness was made in the UK and published in 2006 (140). Twenty-nine RCTs on the effectiveness of adalimumab, etanercept and infliximab as compared to methotrexate were included. Fourteen economic evaluations were additionally included. Different perspectives were used in the included studies. A previously described patient-level simulation model, the Birmingham Rheumatoid Arthritis Model, was further developed and was used to perform the cost-effectiveness analyses. The NICE guidance at the time of the analyses recommended a TNFi as the third DMARD in a sequence of DMARDs which was referred to as the base-case in the cost-effectiveness analyses. The ICER was around £30,000 (€35,669) per QALY in early RA and £50,000 (€59,449) per QALY in late RA when etanercept was used while adalimumab and infliximab yielded higher ICERs. The sensitivity analysis indicated that the results were sensitive to estimates of HAQ progression while treated with TNFi and HAQ progression while treated with sDMARDs, but not to changes in mortality ratios per unit HAQ. The TNFi seemed to be most cost-effective when used last in a sequence including 10 different combinations of sDMARDs. The ICERs for all three TNFi were under £40,000 (€47,559) per QALY when used as last active therapy. When combination therapy was used as first line, *i.e.* MTX plus a TNFi, the ICERs were substantially higher. Use of the societal perspective generated more favourable ICERs (140).

In 2007, a new state-transition microsimulation model from a research group in Sheffield and Manchester was developed with the aim of analysing the cost-effectiveness of TNFi compared to sDMARDs for RA-patients in the UK (141). The modellers used data from the British Society for Rheumatology Biologics Registry (BSRBR), which had three years of follow-up at the time of the analysis. The model simulated individual patients over intervals of six months with a lifetime perspective. HRQoL was assessed both by SF-6D data derived from the SF-36 which was included in the BSRBR and by EQ-5D derived utility which had to be mapped from the HAQ. The model described two clinical pathways; TNFi therapy and ongoing sDMARDs and assessed initial and long-term response to therapy, probabilities of withdrawal of therapy and long-term cost and QALY consequences. At the time of the analysis, current guidelines in the UK recommended that TNFi should be introduced after the failure of at least two sDMARDs. This seemed to be cost-effective, with an incremental cost per QALY of £23,882 (€28,395) and an 84% probability of being cost-effective at £30,000(€35,669). The results were sensitive to assumptions of long-term disability progression, discount rates and choice of utility instrument. The SF-6D generated cost-effectiveness ratios that were double the size of the ratios generated by the EQ-5D (141).

In Sweden, a Markov cohort model was designed in 2010 with the aim of estimating the cost-effectiveness of infliximab use in patients with RA in Swedish clinical practice, based on patient-level data from the Stockholm TNF-alpha follow-up registry (STURE) (142). The model was based the previously described model with five health states based on the HAQ (131) and a subdivision of the states according to DAS28, with a cut-off at 3.2 (131, 136). Two arms were included in the model, one for infliximab therapy and one for natural progression, assuming DMARD treatment in both. The natural progression rate (also assumed for the infliximab arm after infliximab discontinuation), in the base case analysis was an average increase in HAQ of 0.065 per year, a value taken from the UK Early Rheumatoid Arthritis study of 145 patients after a failure of two DMARDs (135). In a sensitivity

analysis, a HAQ-progression rate of 0.031 per year was used, based on a review of HAQ progression from nine published studies and two unpublished data sets, by Scott and co-workers (137). The base case assumptions yielded an ICER of €22,830 including production losses when using EQ-5D utilities estimates from the HAQ and DAS28 (142, 143). When the lower progression rate was used, the resulting ICER was just over €35,000.

A systematic review of 18 selected studies of bDMARDs in the treatment of RA was published in 2011, based on data in the included studies from 2000-2007. CUA was performed in 16 of the 18 included studies, and out of these, 10 studies derived utility weights by transformation of HAQ scores by use of linear regression. The cost-effectiveness of bDMARDs compared to sDMARDs reported results in terms of willingness-to-pay thresholds at Canadian (Can) $50,000 (€37,230) and Can $100,000 (€74,465) per quality-adjusted life year (QALY). At a threshold of Can $50,000, bDMARDs were only cost-effective in one of 35 comparisons in patients who had failed ≥2 DMARDs. Applying a threshold of $100,000, bDMARDs were cost-effective in 14 of 35 comparisons. The ICER values ranged from €33,500 to €456,200. Only one of the studies had a societal perspective; the others had a payer perspective (144, 145).

In the majority of previously published cost-effectiveness studies on TNFi, data from clinical trials with limited follow-up time of maximum two years were used. Exceptions are the study from the UK (141) which used register data with three years of follow-up and the recent Swedish study (142) with almost 10 years of follow-up.

Lack of direct measurement of utility with MAU instruments such as the EQ-5D as well as limited data on long-term treatment effectiveness, resource use and production losses have been limitations in previous studies (144, 146, 147). Thus, no study has assessed the cost-effectiveness using direct measurements of utility within settings of real life long-term data.

# 2 OBJECTIVE

## 2.1 Overall objective

The main aim of this thesis was to compare costs and effectiveness of bDMARDs versus sDMARDs in patients with inflammatory rheumatic joint diseases. Comparing incremental costs to incremental outcomes of alternative interventions contributes to optimal resource allocation within the inflammatory rheumatic joint diseases and the health care sector as a whole.

In the first part of the project, we aimed at exploring two general concepts of patient reported health: PASS and MCII and their relation to MAU instruments.

In the second part, we aimed to investigate the costs incurred by patients with inflammatory rheumatic joint diseases including RA, PsA and AS in Norway.

The third part aimed at performing an economic evaluation of bDMARDs versus sDMARDs using two different MAU instruments. Here, we restricted the analysis to patients with RA.

## 2.2 Objectives by paper

### Paper I

The aim of Paper I was to assess cutpoints for PASS and MCII in MAU instruments and other patient reported outcomes in patients with RA, PsA and AS.

Research questions

- Which cutpoint values in SF-6D and EQ-5D correspond to PASS and MCII in patients with RA, PsA and AS?
- How do the cutpoint values of PASS and MCII in EQ-5D, SF-6D and other patient reported outcomes vary by statistical method used for the assessment?
- How do the cutpoint values vary for PASS and MCII in EQ-5D, SF-6D and other patient reported outcomes in patients with RA, PsA and AS?

### Paper II

The aim of Paper II was to estimate health care costs and production losses during a two-year period after start of DMARDs in patients with RA, PsA and AS.

Research questions

- Do health care costs change during a two-year period after start of sDMARDs/bDMARDs in patients with RA, PsA and AS?
- Do production losses as assessed by the HCA/FCA change during a two-year period after start of sDMARDs/bDMARDs in patients with RA, PsA and AS?
- Do the total costs (including both health care costs and production losses) differ between patients with RA, PsA and AS?

### Paper III

The aims of Paper III were to estimate the additional costs and health benefits of adding a TNFi to a sDMARD, in patients with RA, in a 10-year perspective using data from two observational studies, including both the SF-6D and the EQ-5D.

Research questions

- What are the incremental/additional costs and QALYs of adding a TNFi in combination with sDMARD compared to sDMARD mono- or combination therapy in patients with RA in a 10-year perspective?
- What are the ICERs of adding a TNFi in combination with sDMARD compared to sDMARD mono- or combination therapy in patients with RA in a 10-year perspective using the SF-6D versus using the EQ-5D for assessment of health benefits?

# 3 MATERIALS AND METHODS

The primary data sources for the studies included in this thesis were NOR-DMARD (108) and ORAR (44). Apart from these registries, we applied data from the Drug procurement cooperation (Legemiddelinnkjøpssamarbeid) (LIS), national fee schedules, the Norwegian Directorate of Health,

the Norwegian Medicines Agency, the South-Eastern Norway Regional Health Authority and Statistics Norway.

This section first presents NOR-DMARD, second presents ORAR and third describes which cost items we included, how the quantities of these were assessed and the sources of the cost estimates. Finally, data sources only for the NORA model are presented.

## 3.1 Data

### 3.1.1 The NOR-DMARD study (Papers I-III)

Patients with inflammatory joint diseases were included in NOR-DMARD when they started treatment with sDMARDs or bDMARDs regimens. The patients were recruited at five different rheumatology departments in Norway: Oslo, Drammen, Lillehammer, Trondheim and Tromsø. These departments cover a total catchment area of about 1.4 million inhabitants. The register started in December 2000, patients have been followed longitudinally since the start and new cases have continuously been included (108, 148, 149). The attending rheumatologist initiated the DMARD regimen based on clinical judgement and existing treatment recommendations. Clinical assessments were performed at baseline, after 3, 6 and 12 months and then yearly up to change of DMARD treatment or treatment termination (108, 149). NOR-DMARD is consequently a prospective multicentre longitudinal observational study (149). A new baseline registration was performed each time a patient started a (new) DMARD-regimen. This means that each patient could be registered twice or more if the treatment was changed. In February 2012, NOR-DMARD comprised 11,875 treatment courses initiated in 7,675 individual patients.

The NOR-DMARD database includes information on diagnosis, age, sex, type of medication, disease duration, education and employment status as well as outcome measures for disease activity and generic measures of health. The generic instruments included are the SF-36 (109), which can be used for the calculation of the SF-6D (110), and the EQ-5D (111). In this thesis, EQ-5D is used to refer to the 3L-version. In addition, NOR-DMARD includes data on the patients' use of health care services (150).[1]

NOR-DMARD comprises four data files; one including demographic variables, outcome measures, variables for use of health care services and level of participation in the workforce. Visits to GPs, outpatient clinics, physical therapists, private rheumatologists as well as imaging examinations are registered. Hospital- and rehabilitation stays are included. Level of participation in the workforce is registered at six levels; retired, on disability pension or rehabilitation benefits, on sick leave, fully employed, working part time and other (students, unemployed, etc.). The second file, "DMARD Cort", includes details about DMARD use. The type, frequency (daily, weekly, monthly or other) and doses (in milligram or gram) for each DMARD and patient are registered in this file. The third file, the "Infusion log" includes information on dates and doses of drugs taken as infusions. Finally, the

[1] The study was approved by the Norwegian Data Inspectorate and the Regional Ethics Committee (REK) of South-Eastern Norway. Patients gave written informed consent according to the Declaration of Helsinki before participation.

"ConMed" file includes all concomitant medications, registered by start- and stop dates and the frequency are either registered as a regular regimen or taken as needed.

For Paper I, the August 2008 version of NOR-DMARD files was applied, including 8,078 treatment courses and 5,417 unique individuals. In Paper II, the February 2010 updated version was applied, encompassing 9,919 treatment courses and 6,518 individual patients. In Paper III, the February 2012 version was used, including 11,875 treatment courses in 7,675 individual patients.

The different formats used for the registration of when and how much drugs patients used in the different data files were recalculated to periods. In Paper II, we assessed the use of resources in six-month periods. In Paper III, we considered it better to perform three-month period assessments. The first two registrations in NOR-DMARD have a three-month interval between assessments and we wanted to take advantage of all possible information in the observational data for the population of the Markov model.

### 3.1.2 The ORAR study (Paper III)

ORAR was established in 1994. Adult patients in Norway's capital city Oslo (population approximately 450,000 when ORAR was established) with an RA diagnosis were included in the register (44). Five mail surveys have been performed with comprehensive questionnaires during the period 1994-2009 (151). The register was continuously updated with new cases, and patients were withdrawn at death or when they moved outside Oslo (44). The patients in the study might have stayed on treatment/no treatment, changed treatment or discontinued any treatment. In 1994 no patients used TNF inhibitors while in 2001 3.1 % were on TNF inhibitors (152). sDMARDs and prednisolone were in the whole ORAR used respectively by 36 % and 41% of patients in 1994 and 48% and 43% in 2001 (151, 152). We used SF-6D data from ORAR (1994-2001) to assess the development in HRQoL for RA-patients on sDMARD treatment. In total, 412 patients had SF-36 data (and consequently SF-6D utilities could be estimated) in 1994 and in 2001. Of those, 173 were on treatment with a sDMARD and 87 used MTX in 1994. The information on changes in HRQoL over a seven-year period from ORAR was used for informing the development in HRQoL for patients in the synthetic-strategy in the NORA model.

### 3.1.3 Resource use and pricing (Paper II-III)

We identified the relevant types of resource use, assessed the quantities and multiplied these with the unit costs. For unit prices of pharmaceuticals, we used the Norwegian Medicines Agency's web site (2010 Paper II and 2012 Paper III) and the Physician's Desk Book from 2004 for older drugs (conventional DMARDs, glucocorticoids, anti-inflammatory agents and analgesics). For over-the counter (OTC) analgesics we used the prices from a local pharmacy. The prices of biologics were taken from the Drug procurement cooperation (Legemiddelinnkjøpssamarbeid) (LIS) (http://www.lisnorway.no). This organization is the centre for contracts for the purchase of pharmaceuticals for the health authorities in Norway.

The cost of taking a drug by infusion was calculated by a DRG cost weight (DRG 808H) according to the activity-based funding from the Norwegian Directorate of Health (153). The cost of a hospital stay, either at a general or rheumatology department, was based on the activity based-funding scheme (DRG 242B and DRG 242C) as was the cost of an outpatient visit (DRG 908C). The cost of a visit to a general practitioner was based on the tariff of the Norwegian Medical Association (154) and the South-Eastern Norway Regional Health Authority. The cost of a visit to a private practising

rheumatologist was divided in a 1st comprehensive visit and following visits, both of which were taken from the Norwegian Medical Association and the South-Eastern Norway Regional Health Authority. The cost of taking an image was taken from the Norwegian Directorate of Health (Regulations for financing outpatient radiology 2012)(155). Visits to physiotherapists were also divided in a 1st comprehensive examination and following visits. The physiotherapy costs were assessed according to the tariff from the Norwegian Physiotherapist Association. The cost of a rehabilitation stay was taken from an average stay for a patient with RA at Skogli rehabilitation centre in Norway. Prices were exclusive of value-added tax. Patient co-payments and travel costs were added to all cost items, except for the drug costs and the imaging examination. The imaging examination did not include a travel cost since this examination was supposed to take place at the same time as another visit/examination in most cases.

The cost of lost productivity was estimated as the yearly pre-tax income using data from Statistics Norway (2010 Paper II and 2012 Paper III) (NOK 470,900 in 2012). We added 40% to the wage rate to account for the social cost of labour. Patients who were in the two categories: "retired due to age" or "other" (students, unemployed, etc.) were excluded from the productivity loss calculations, as they were assumed to have no productivity losses. For patients who were unable to work, including patients on disability pension, rehabilitation and sick leave, the mean income plus the social cost of labour was added as a productivity loss. We assumed full time work as the normal workforce participation. For patients who were able to work part time, this productivity cost was reduced in proportion to the time worked. We adopted the HCA in Paper II and III, assuming that the productivity loss persisted throughout the analysed time. For comparison, we calculated the production losses with the FCA, assuming a friction period of 5 months, in Paper II. A friction period is the assumed time for an absent worker to be replaced by another.

### 3.1.4 Data only for the NORA model (Paper III)

In the NORA model, a start cost was added to all patients in the TNFi-strategy, including obligatory examinations before the initiation of a bDMARD therapy. This start cost encompassed the cost of taking a chest x-ray and of testing for hepatitis B+C and tuberculosis. Further details of the cost calculations to the model are presented in online supplementary text in Paper III. We used mortality tables from Statistics Norway, adjusted for the increased mortality in RA, for the estimation of probability of death.

### 3.1.5 Summary of data sources by paper

Table 11. Summary of data sources by paper

| Paper | Data source |
|---|---|
| Paper I | The main file in NOR-DMARD |
| Paper II | The main file in NOR-DMARD<br>The three files detailing medication use in NOR-DMARD<br>LIS<br>National tariffs<br>The Norwegian Directorate of Health<br>The Norwegian Medicines Agency<br>The South-Eastern Norway Regional Health Authority |
| Paper III | The main file in NOR-DMARD<br>The three files detailing medication use in NOR-DMARD<br>The ORAR<br>LIS<br>National tariffs<br>The Norwegian Directorate of Health<br>The Norwegian Medicines Agency<br>The South-Eastern Norway Regional Health Authority<br>Statistics Norway |

## 3.2 Statistics and methods

This section presents the statistical methods applied by each paper. Under Paper III, the development of the NORA model is described.

### 3.2.1 Paper I

Descriptive statistics including mean and standard deviation (SD) were used to describe population characteristics.

The question included in the NOR-DMARD questionnaire regarding PASS was: "Is your current condition satisfactory, when you take your general functioning and your current pain into consideration?" The response options were "yes" or "no". Regarding MCII, there were two questions. The first question was "Since you started treatment in this follow-up study, is your health condition improved, unchanged or worse?" The answers were measured on a 5-point Likert scale, from "much better" to "much worse". The two alternatives: "much better" and "better" were merged and regarded as a considerable improvement in the analyses. The second question was: "Has the treatment in this follow-up study improved your health condition considerably?" and the response options were "yes" or "no" (108).

For assessing levels/changes in patient reported outcomes that correspond to PASS/MCII, we used two methodologies. These were the 75$^{th}$ percentile approach and the 80% specificity approach using receiver-operating characteristic (ROC) analyses (103-105). In sensitivity analysis we identified cut offs that yielded the smallest number of both false positive and false negatives, *i.e.* the maximum accuracy method (103, 104).

The 75th percentile approach includes patients who have responded positively to the PASS/MCII question. In patients who have responded positively, the 75th percentile of the cumulative distribution of scores in each examined outcome measure is assessed as the cut off value. When using PASS and MCII this corresponds to patients who consider their health state as satisfactory or have experienced an important improvement, respectively (156).

ROC analysis is a method for visualizing, organising and selecting classifiers. A classifier in this case is a mapping from instances to predicted classes. Given a classifier and a set of instances, a contingency table can be constructed to illustrate possible outcomes. These classifications are used in the ROC analysis.

Table 12. A classification system for true and hypothesized classes. Adapted from: Pattern Recognit Lett. 2006;27(8):861-74.

| | | True class |
|---|---|---|
| Hypothesized class | **T**rue **P**ositives (**TP**) | **F**alse **P**ositives (**FP**) |
| | **F**alse Negatives (**FN**) | **T**rue Negatives (**TN**) |
| | **P** | **N** |

The sensitivity (true positive rate (TP rate)) is defined as the probability that the test is positive among those with the disease/condition, true positives (TP)/total positives (P).

The specificity is defined as the probability that the test is negative among those without the disease/condition; true negatives (TN)/total negatives (N). This is the same as 1-the false positive rate (FP rate) (157).

In addition to the 75th percentile approach for assessing a classifier, there is the 80% specificity method. We chose this as our primary approach, based on previous studies (106). The 80% specificity method includes patients that have responded negatively to the PASS/MCII question as the start point for the analysis. This means that 80% of patients who state that they are not in a satisfactory condition or have not experienced an important improvement have a score on the measurement in question that correspond to the cut off value or a worse health state/less improvement.

We assessed the area under the curve (AUC) by the ROC curve analysis for PASS and MCII and the examined outcome measures. In sensitivity analysis, we used the maximum accuracy method. The cut off value using this approach is the largest numeric sum of sensitivity and specificity (103, 104). Sensitivity and specificity can be traded against each other, the maximum accuracy can possibly be identified with several combinations of sensitivity and specificity, and the choice of cut off value is then equivocal.

MCII is a rating of the extent to which patients' health conditions have improved, and this type of rating may be called a transition rating. The MCII was tested for validity according to a method for transition ratings suggested by Guyatt et al (107). Correlation was used to examine the extent to which MCII was associated with changes in other patient-reported instruments (EQ-5D, SF-6D, Patient global VAS, Pain VAS and MHAQ), from treatment initiation to the assessment after 3 months. In addition, changes in the different dimensions in EQ-5D were examined by correlation to assess the association to the MCII.

### 3.2.2 Paper II

Descriptive statistics including mean and range were used for population characteristics.

Missing information on dose frequencies for each particular DMARD was replaced by the corresponding mode frequency for the DMARD in question. For infusion drugs, we used irregular schemes in accordance with normal drug distributions for each drug. These are often distributed with smaller intervals between infusions in the start of a treatment period and larger intervals after a start-up period. Missing information on doses of DMARDs was replaced by multiple imputations using multivariate normal and linear regression models. These were imputed separately for each DMARD. Multivariate normal regression models replaced missing values for other cost variables, such as GP visits and hospital stays.

Differences in costs between first and last six-month periods were normally distributed and tested by paired *t*-tests with $p<0.05$ considered significant.

Confidence intervals of mean total costs over the two-year analysed period were calculated by 10,000 bootstrapped estimates since the total costs were not normally distributed.

Differences in median total costs across diagnoses were tested by Kruskal-Wallis equality-of-populations rank test. CIs for median differences were calculated by the Hodges-Lehmann's estimator.

Data were analysed using Stata MP11, College Station, TX, USA.

### 3.2.3 Paper III – the development of the NORA model

We developed a Markov cohort model; the NORA model, with the aim of performing a cost-utility analysis of treatment options with DMARDs for RA patients. An overview of the model and an explanation of the development is given in the following and further descriptions are presented in Paper III with Supplementary Data. We included two strategies, treatment with sDMARDs in combination with a TNFi (TNFi-strategy) and treatment with sDMARDs alone (synthetic-strategy).

In the model, patients are in different states that are mutually exclusive and collectively exhaustive. Both strategies include seven main states, based on SF-6D utility levels (0.296-0.4, >0.4-0.5, etc. up to 1.0), in addition to death. We used a cycle length of 3 months (a quarter of a year), and the time perspective was 10 years. In the TNFi-strategy, patients started with a TNFi (adalimumab, certolizumab, etanercept, golimumab or infliximab) combined with a sDMARD. Patients could remain on the same therapy, switch TNFi, switch to mono TNFi treatment or switch to another bDMARD (abatacept, rituximab or tocilizumab). Patients on all types of bDMARD treatment could be in any of the seven health states and move between them once per quarter. Patients could also move to no DMARD (a specific state in the model), in which they stayed for one year before they moved back to bDMARD treatment. Additionally, patients could move to sDMARDs. Here, they remained in the state for the rest of the analysed period. Finally, patients might die.

In the synthetic-strategy, patients were on sDMARD treatment and did not switch to bDMARD. In line with the TNFi-strategy, patients in the synthetic-strategy could be in any of the seven health states and switch between them and they might die.

We developed the model with health states based on the SF-6D. The utility values in the states were the mean values for patients in the state at start of treatment. Mean values of both SF-6D scores and EQ-5D scores were used in sensitivity analyses.

For the TNFi-strategy we had comprehensive and detailed data on HRQoL from NOR-DMARD. For the synthetic-strategy the data from NOR-DMARD were not considered to be suitable since patients on sDMARD were on average in better health states than patients on TNFi therapy. Thus, we needed other/supplementary data for the population of the synthetic-strategy. We decided to use SF-6D data from ORAR, derived from patients with routine treatment before bDMARDs came into common use. We wanted to find the subgroup of patients from ORAR that best matched the patient population from NOR-DMARD who started with a TNFi after failure of on average 2.1 sDMARDs.

The ORAR questionnaire was sent to all RA patients in the region of Oslo. 452 patients with SF-6D data in 1994, and 412 patients with SF-6D data also in 2001 were available for the analysis of development in utility over time. Mean SF-6D were 0.645 (CI:0.632; 0.657) in 1994 and 0.625 (CI:0.611; 0.638) in 2001. 173 patients were on sDMARD treatment in 1994 and had SF-6D data both in 1994 and in 2001. Their mean SF-6D was 0.631 (CI:0.613; 0.649) in 1994 and 0.634 (CI:0.614; 0.655) in 2001, thus no real change in in utility scores. The subgroup of patients who used MTX in 1994 (n=87) had a mean SF-6D of 0.610 both in 1994 (CI:0.586; 0.633) and in 2001 (CI:0.585; 0.633). We chose to use the 173 patients who used sDMARD in 1994 as the basis for the synthetic-strategy in the model. The distribution of SF-6D scores at start in NOR-DMARD and ORAR is presented in table 13 and figures 5a-b.

Table 13. Distribution of SF-6D among RA patients at initiation of treatment with sDMARDs plus TNFi (TNFi-strategy, NOR-DMARD) and the distribution of SF-6D among RA patients from ORAR in 1994 for all patients, patients on sDMARD treatment in 1994 and patients on MTX treatment in 1994.

| Health states | NOR-DMARD n=810 | | ORAR all n=412 | | ORAR DMARD n=173 | | ORAR MTX n=87 | |
|---|---|---|---|---|---|---|---|---|
| Range 0.296-1.0 | Frequency | % | Frequency | % | Frequency | % | Frequency | % |
| 0.296 ≤ 0.4 | 51 | 6 | 7 | 2 | 3 | 2 | 2 | 2 |
| >0.4 ≤ 0.5 | 109 | 13 | 36 | 9 | 14 | 8 | 12 | 14 |
| >0.5 ≤ 0.6 | 312 | 39 | 139 | 34 | 66 | 38 | 33 | 38 |
| >0.6 ≤ 0.7 | 193 | 24 | 101 | 25 | 45 | 26 | 22 | 25 |
| >0.7 ≤ 0.8 | 84 | 10 | 67 | 16 | 28 | 16 | 12 | 14 |
| >0.7 ≤ 0.9 | 57 | 7 | 52 | 13 | 13 | 8 | 6 | 7 |
| >0.9 ≤ 1.0 | 4 | 0 | 10 | 2 | 4 | 2 | 0 | 0 |
| Total | 810 | | 412 | | 173 | | 87 | |

Sources: ORAR and NOR-DMARD

Figure 5a. Distribution of SF-6D scores among RA-patients at initiation of TNF inhibitor plus synthetic DMARD treatment ((TNFi-strategy, NOR-DMARD (n=810)).

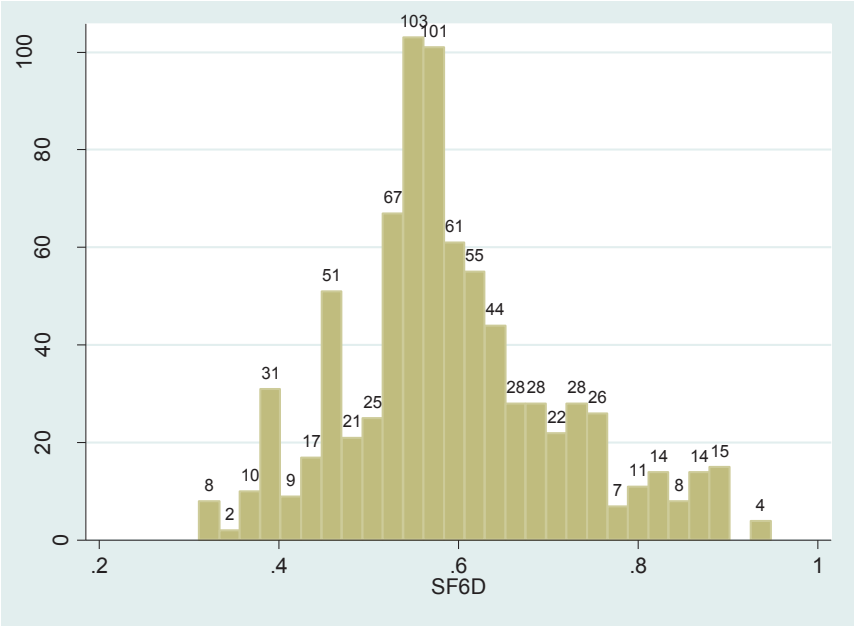
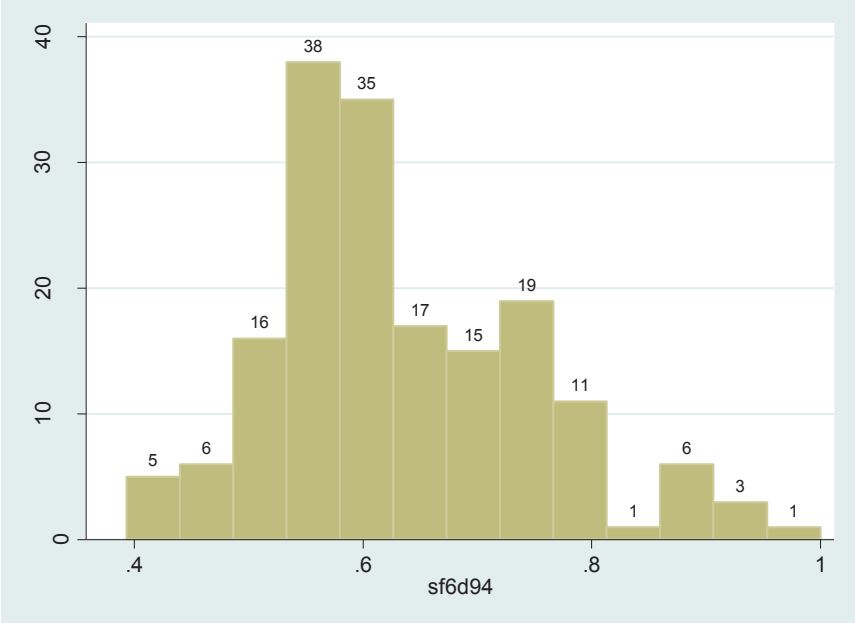



Figure 5b. Distribution of SF-6D scores among RA-patients on synthetic DMARD treatment. Data from ORAR (n=173).



The cost estimates in the TNFi-strategy are based directly on the costs from patients starting treatment with a TNFi+sDMARD *i.e.* their resource use from treatment start to end of follow-up. For the cost calculations see section 3.1.3 Resource use and pricing (Paper II-III) and section 3.1.4 Data for the NORA model (Paper III). A difference in the cost calculations in Paper II and the calculations in Paper III is the division of resource use in six-month versus three-month periods. The cost estimates for the synthetic-strategy were based on the resource use of the patients who later started treatment with a TNFi+sDMARD but three months before treatment start when they still were on sDMARD treatment. Missing information on the use of DMARDs in NOR-DMARD was replaced by multiple imputations using multivariate normal and linear regression models both in Paper II and Paper III. Missing values for other cost variables, such as GP visits and hospital stays were replaced by zero in Paper III, assuming that a missing value most probably meant that no visit/stay/examination had taken place.

The transition probabilities for patients in the TNFi-strategy are based on data in NOR-DMARD from patients starting treatment with a TNFi + sDMARD. The probabilities of a change of health state during the first two quarters, *i.e.* the probability of a transition from a health state to all other health states from the first to the second quarter and from the second to the third quarter were calculated separately since we had quarterly updates from the register for these transitions. Thereafter, the yearly probabilities of a change, *i.e.* from quarter 4, 8, 12, 16, 20, 24, 28, 32, 36 and 40 were calculated separately since the probability of a change was higher when the yearly assessments were scheduled. The probabilities of changing between the other quarters were estimated separately since these probabilities were lower than the yearly changes. The probabilities of a switch to sDMARDs and to no DMARD were also calculated by ratios directly from the data in NOR-DMARD. In sensitivity analyses we explored another way of estimating the probabilities over time. We grouped the transition probabilities in: 1-2 quarter, 2-3 quarter, 3-9 quarter (6 months-2 years), 9-25 quarter (year 3,4,5,6) and 25-41 quarter (year 7,8,9,10). The result in QALYs over 10 years with the alternative approach was very close to the base case analysis. The base case analysis yielded 6.54 and the alternative approach yielded 6.53 QALYs. In the base case analysis we had four groups of transition probabilities and in the sensitivity analysis five. Thus, we chose the alternative with fewest different transition probabilities. A transition probability in the synthetic-strategy was the same as the initial probability of being in the particular state. This might seem odd but was because patients were supposed to end up at the same average health utility level after 10 years as at the initiation of the model, based on data from ORAR.

We performed 1,000 Monte Carlo simulations to assess long-term consequences in terms of health outcomes and costs when using bDMARDs in addition to sDMARDs versus sDMARDs alone. At start of the simulation, the cohort was distributed across health states according to findings in NOR-DMARD (n=810).

We performed deterministic sensitivity analyses of using SF-6D utility versus EQ-5D utility scores as the basis for the QALY calculations in the NORA model. Further, we performed deterministic sensitivity analysis including health care costs only and health care costs plus productivity losses and undiscounted/discounted costs and QALYs. Probabilistic sensitivity analyses were performed for all costs and QALY estimates as well as for transition probabilities between health states. In NORA, we used Dirichlet distributions when estimating the probability of changing health state (model states) from one time period to another. We used gamma distributions for health care costs, total costs

(including production losses) and utility decrements/disutilities. A total of 78 different probabilistic distributions were specified and included in the model. The specific mean values and standard errors for each distribution are presented in online supplementary tables in Paper III.

The uncertainty in the results was presented in cost-effectiveness planes and in cost-effectiveness acceptability curves for four different combinations of costs and QALYs with discounted values. The base care analysis included health care costs and EQ-5D utilities. The base case willingness-to-pay (WTP) threshold was set at €67,300 (NOK 500,000). NOK 500,000 in 2005 value has been proposed as the value of a QALY consistent with a value of a statistical life with full health of 15 million, for use in economic analyses covering different societal sectors (158). An update of €67,300 (NOK 500,000) in 2005 value would be €79,100 (NOK 588,000) in 2012 value. However, in recent Norwegian guidelines on economic evaluations in the health care sector, no specific WTP threshold value was given. The recommendation is to perform the analysis with a range of WTP-thresholds (25). We applied WTP-thresholds from NOK 400,000 (€53,800) to NOK 700,000 (€94,200). Finally, we estimated the EVPI from the Monte Carlo simulations for the base case analysis.

Data were analysed using Stata MP11, College Station, TX, USA and the model was developed in TreeAge Pro 2012 and 2013.

# 4 SUMMARY OF RESULTS AND MAIN FINDINGS

## 4.1 Paper I

Research questions: Which cutpoint values in SF-6D and EQ-5D correspond to PASS and MCII in patients with RA, PsA and AS?
How do the cutpoint values of PASS and MCII in EQ-5D, SF-6D and other patient reported outcomes vary by statistical method used for the assessment?
How do the cutpoint values vary between patients with RA, PsA and AS?

### 4.1.1 Results

The cutpoints for PASS varied from 0.62 to 0.69 in EQ-5D and from 0.60 to 0.64 in SF-6D with the 75$^{th}$ percentile approach, depending on diagnosis. The estimated cutpoints using the 80% specificity method varied from 0.69 to 0.73 in EQ-5D and from 0.65 to 0.66 in SF-6D. Thus, the estimated cutpoints varied somewhat with diagnosis and more with the methodology used. This means that among patients who stated that they had a satisfactory health state, 75% had an EQ-5D score of approximately 0.65 or higher and 75% had an SF-6D score of 0.60 or higher. In patients who stated that they were not in a satisfactory health condition 80% had a score of approximately 0.70 or lower in EQ-5D and 0.65 or lower in SF-6D. The cutpoints in EQ-5D seem to be slightly higher than the corresponding values in SF-6D. Both methodologies yielded cutpoints in the MAU instruments that are far from perfect health (1.0). PASS cutpoints for the health status measures were similar across diagnoses.

The cutpoints for MCII varied from 0 (RA and PsA) to 0.04 (AS) in EQ-5D and from 0.01 (RA) to 0.05 (AS) in SF-6D with the 75$^{th}$ percentile approach, depending on diagnosis. Using the 80% specificity method, the cutpoints varied from 0.10 (RA) to 0.19 (AS) in EQ-5D and from 0.07 (PsA) to 0.09 (AS) in SF-6D. The cutpoints varied thus quite extensively by methodology used. For EQ-5D, we found the largest difference between diagnoses with cutpoints varying from 0.10 to 0.19 using the same

methodological approach, whereas MCII cutpoints for SF-6D, MHAQ, pain VAS and patient global VAS were rather similar across diagnoses.

The area under the curve (AUC) assessed by the ROC curve analysis for the MCII yielded estimates from 0.68 to 0.83 with different instruments. The AUC for the MCII versus EQ-5D and SF-6D was low (0.68-0.77) which indicates a low agreement between the instruments.

The MCII resulted in an overall inferior agreement compared to the PASS that yielded estimates from 0.75 to 0.85 across instruments and diagnoses.

### 4.1.2 Main findings

The results from our study revealed that estimated cut off values for classifying a patient to be in a satisfactory health condition or as having experienced a minimal clinical important improvement vary by statistical methodology used for the assessment. The 80% specificity approach is a more conservative approach than the 75$^{th}$ percentile approach, which means that patients have to have better health and larger numerical improvements as assed by other outcome measures for being classified to be in PASS and having experienced a MCII. Both methodologies yielded cutpoints for PASS in the MAU instruments that are far from perfect health.

## 4.2 Paper II

Research questions: Do health care costs/production losses change during a two-year period after start of sDMARDs/bDMARDs in patients with RA, PsA and AS?
Do the total costs (including both health care costs and production losses) differ between patients with RA, PsA and AS?

### 4.2.1 Results

The largest cost component across all diagnoses and treatment types was production losses, followed by the cost of DMARDs for biologic treatment and the cost of in-hospital care for synthetic treatment. In the base case analysis, we used the HCA, assuming that the production loss persisted throughout the analysed period (two years). In sensitivity analyses, we applied the FCA, assuming a friction period of five months.

We compared costs over a two-year period for patients with RA, PsA and AS after start of DMARDs. From the first six-month period to the last six-month period (defined as the second-year cost divided by two) on DMARD treatment, the total costs declined significantly in all diagnoses for bDMARD treatment ($p$<0.05). Changes in total costs for patients on sDMARD treatment were significant for RA and PsA but not for AS. Both health care costs and production losses declined but not all combinations of diagnoses and therapy declined significantly. When the FCA was used, total costs and production losses declined significantly in all combinations of diagnoses and therapy, as expected using this approach.

The total two-year costs were similar across diagnoses for patients on sDMARD treatment. For bDMARD treatment, RA patients had the highest mean and median total costs (mean €121,900) (NOK 950,800) followed by AS (mean €115,300) (NOK 900,000)and PsA (€111,200) (NOK 867,300) but the differences were not statistically significant, either using the HC or the FCA.

The annual health care costs for RA patients on sDMARD treatment were approximately €3,400 (NOK 26,300) and for patients on bDMARD treatment the corresponding costs were €19,600 (NOK

152,600). The annual costs including production losses (HC approach) for RA patients on sDMARD treatment were approximately €32,200 (NOK 250,900) and for patients on bDMARD treatment the corresponding costs were €60,900 (NOK 475,400).The corresponding costs were somewhat smaller for PsA patients and slightly higher for AS patients.

### 4.2.2 Main findings

Biologic DMARD treatment entails considerable drug cost, but the total costs (including health care costs and production losses) decline during the first two years on treatment in both RA, PsA and AS. The total costs are similar across RA, PsA and AS.

## 4.3 Paper III

Research questions: What are the incremental/additional costs and QALYs of adding a TNFi in combination with sDMARD compared to sDMARD mono- or combination therapy in patients with RA in a 10-year perspective?
What are the ICERs of adding a TNFi in combination with sDMARD compared to sDMARD mono- or combination therapy in patients with RA in a 10-year perspective using the SF-6D versus the EQ-5D?

### 4.3.1 Results

The discounted health care costs were €65,584 over the 10-year period in the synthetic-strategy and €124,942 in the TNFi-strategy (€436,517 and €475,266 respectively when production losses were included).

The synthetic-strategy generated 4.78 and 3.82 QALYs in total over a 10-year period with SF-6D and EQ-5D, respectively. The TNFi-strategy generated 5.42 QALYs with SF-6D and 4.79 QALYs with EQ-5D.

The incremental cost per QALY was €92,557 (NOK 687,697) using SF-6D and €61,285 (NOK 455,351) using EQ-5D, excluding production losses. Including production losses, the ICERs diminished to €60,227 (NOK 447,488) using SF-6D and €39,841 (NOK 296,019) using EQ-5D.

Including health care costs and using SF-6D,there was zero probability that TNFi treatment was cost-effective at a willingness-to-pay (WTP) threshold of €67,300 (NOK 500,000). If EQ-5D was used, the probability that TNFi-treatment was cost-effective was 90% at the same WTP-threshold. When also production losses were included, TNFi-treatment was cost-effective, irrespective of the type of MAU instrument. Using SF-6D, the probability that TNFi-treatment was cost-effective was 89%. EQ-5D utilities yielded a probability of 100% for the same strategy, already at a WTP threshold of €53,800 (NOK 400,000).

The base care analysis included health care costs and EQ-5D utilities. The EVPI for the base case scenario was NOK 1,494 (€201) per patient. Since the cost-effectiveness acceptability curve indicated a 10% possibility of sDMARDs to be the most cost-effective choice when TNFi was the most cost-effective choice in the majority of simulations (90%), there is a cost of making the wrong decision in 10% of cases. In contrast, for the scenario including production losses and EQ-5D utilities, there is no uncertainty of the cost-effectiveness of TNFi-therapy given a WTP threshold of €67,300 (NOK 500,000) and the value of additional information is zero.

### 4.3.2 Main findings

TNFi-treatment for RA is cost-effective when accounting for production losses, assuming a Norwegian willingness-to-pay level of approximately €67,300. Excluding production losses, TNFi-treatment is cost-effective using EQ-5D, but not SF-6D.

# 5 DISCUSSION

This section will critically evaluate the data and methods applied and discuss the main findings of the thesis. I will start by discussing the data followed by a methodological discussion. Here, the characteristics of the NORA model is the main issue. Thereafter I will discuss the main findings in relation to other studies.

## 5.1 Discussion of data

NOR-DMARD is the main data source for the thesis and the data discussion is limited to the NOR-DMARD register. Some issues related to the use of data from ORAR are included in the discussion of methods, in the following section.

### 5.1.1 NOR-DMARD

The NOR-DMARD longitudinal observational study of routine care for patients with inflammatory rheumatic joint diseases includes consecutive patients when they start treatment with sDMARDs or bDMARDs. The choice of treatment is made by the rheumatologist in cooperation with the patient, and there is no random allocation to type of treatment. The lack of randomization entails both strengths and weaknesses. The main strength is that the whole range of patients taking a medication can be included. The external validity, which concerns to what degree we can draw general conclusions from study results, is likely to be good when we use data from extensive observational studies. When using data from RCTs, there is a risk of obtaining study results that can only be extrapolated to selected patient groups. However, the hospital areas included in NOR-DMARD do not cover all inhabitants in Norway but approximately 30% of the population. Both the capital area of Oslo and rural areas around the towns of Lillehammer and Tromsø are included, reflecting different areas of the country. Thus, the patients are likely to be representative of treatment in Norway.

An additional strength is the long follow-up time. We had data from December 2000 to February 2012 but due to few observations with follow-up over 10 years we restricted the analysis to a follow-up period of 10 years. However, 10 years is a longer follow-up period than normally is scheduled in RCTs. A weakness in NOR-DMARD is that patients undertaking different treatment options are not directly comparable, as in a RCT. Initial analyses showed that RA patients initiating MTX monotherapy were in an overall better health state as measured by DAS28, MHAQ, EQ-5D and SF-6D than those initiating TNFi + MTX therapy.

The data collection in NOR-DMARD is comprehensive compared to most other international real life studies, and the patients are assessed at fixed pre-defined time points. A research nurse or physician collected demographic information and details about the disease status, such as joint counts and laboratory findings. Changes in health and working status and DMARD dosages were registered by the nurse or physician. The patients were asked about their use of health care during the period from previous to current visit. Patients filled in a written questionnaire for patient-reported outcomes, including both MHAQ, EQ-5D (from 2006) and SF-36. The resource utilization was based on self-report from patients supported by communication with the study nurse. This use of self-report

entails a risk of both under- and/or over-reporting if patients forgot for example number of visits to the physical therapist or forgot the date when he/she were on part time work/sick leave. Information from other registry sources would have given more reliable information as for example in the recent Swedish analyses of sick-leave and disability pension data related to disease onset, initiation of treatment and most recently also the SWEFOT study (159-161).

## 5.2 Discussion of methods

The NORA model is the main tool of the thesis and in this section; I discuss the development of the model. First, I discuss the included treatment strategies followed by a note on another possible strategy. Further, I discuss the health states and the transition probabilities in the model. Some adjustments of data management were made between Paper II and III and a discussion of these differences is available at the end of the methodological discussion.

### 5.2.1 The development of the NORA model

Current treatment recommendations suggest the use of bDMARDs for RA-patients responding insufficiently to MTX and/or other sDMARD strategies (14). The model was developed to compare costs and outcomes for sDMARDs in combination with a TNFi versus treatment with sDMARDs alone (mono- or combination therapy). All included treatment combinations are described in supplementary material to paper III. Since the costs of the bDMARDs are significantly higher, the cost-effectiveness of treatment with bDMARDs versus sDMARDs is relevant to assess in a Norwegian setting. The comparison to sDMARDs is most relevant as recent research suggest that treatment with combinations of sDMARDs can have comparable effectiveness to the combination of TNFi and MTX in terms of disease activity and/or HRQoL (162-166).

RCTs on the effectiveness of triple therapy, *i.e*. the combination of MTX, sulfasalazine (SSZ) and hydroxychloroquine (HCQ), compared to MTX and TNFi, have indicated that outcome in terms of disease activity is similar between strategies in trials with follow-up of up to 2 years (163, 164, 167). The SWEdish FarmacOTherapy (SWEFOT) RCT showed significantly better response as measured by the EULAR response criteria in patients treated with MTX plus infliximab compared to triple therapy at 12 months (168) and significantly better radiographical results after 2 years but no significant differences in the EULAR response criteria or in QALY gain after 2 years (161, 167). The Treatment of Early Aggressive Rheumatoid Arthritis (TEAR) study reported similarities in clinical outcomes but radiographic differences in favour of MTX plus etanercept compared to triple therapy after 2 years (163). However, the radiographic results are not consistent between studies since improved radiographic outcome for patients with MTX plus etanercept were lacking compared to triple therapy in the recent RA Comparison of Active Therapies (RACAT) study with a follow-up of 48 weeks (164). In addition, the Finnish NEO-RACO trial reported similar clinical and radiographic results in patients with triple therapy and prednisolone with or without initial infliximab for 6 months. The initial infliximab group had better radiographic results after 2 years but the differences were not statistically significant at 5 years of follow-up (165).

Additionally the inclusion of glucocorticoids in a treatment strategy seem to matter in achieving reduced disease activity (169-172). Already in 1997, results from the COBRA trial (Combinatietherapie Bij Reumatoide Arthritis) demonstrated improved disease control in a combination therapy including prednisolone, MTX and SSZ in comparison to SSZ alone (169). Similar clinical results could be reproduced with lower doses of prednisolone and MTX (171). A Finnish study

revealed the superiority of triple therapy combined with prednisolone in inducing remission, compared to single DMARD therapy with or without prednisolone (172). Recent analyses from the Treatment in the Rotterdam Early Arthritis Cohort study (tREACH) report that initial triple therapy in combination with glucocorticoids gave faster response in achieving reduction in DAS- and HAQ-scores compared to initial MTX and glucocorticoids. No differences in radiographic progression were detected between the strategies and there were no differences in distributional method of glucocorticoids, *i.e*. oral or a single intramuscular injection (173).

To sum up, the above findings indicate that combinations of sDMARDs, including glucocorticoids seem to be effective in achieving remission at least in studies up to 2 years. However, the long-term effects of different treatment options of disease activity, radiographic changes and not least in terms of HRQoL are unclear. A cost-effectiveness comparison between bDMARD regimens and combination of sDMARD regimens in NOR-DMARD would have been interesting but the number of patients with sDMARD combination therapies is rather limited and the combination of sDMARDs is heterogeneous (149).

We had access to NOR-DMARD and ORAR and these were obvious data sources since no other Norwegian data sources for development in HRQoL and transition probabilities were available. We considered that use of data sources from other countries was not relevant due to high quality Norwegian registers and the unique access to longitudinal data on utility measures. Data from these registries were used for estimation of costs, QALYs and not least the transition probabilities between health states.

#### *5.2.1.1 Health states*

To our knowledge, our RA model is the first in which the health states were defined based on HRQoL data. Because both of our observational studies; NOR-DMARD and ORAR, included the SF-36 for calculation of SF-6D utilities, we could base the health states directly on these utility data. Previous health economic models have used HAQ-score, DAS28 or other disease specific instruments; requiring translation of the scores from the instrument in question into utilities. Lillegraven et al fitted linear regression models for the estimation of EQ-5D and SF-6D scores in RA-patients including HAQ among the independent variables. HAQ did only explain about 10-15% of the variance in the regression models and the authors question the validity in estimating utility values from HAQ (174).

In initial analyses we tested distributions of utility scores (SF-6D and EQ-5D) and costs in health states based on the MHAQ, DAS28, EQ-5D and SF-6D. The SF-6D gave the smoothest distribution of patients in the states, the costs increased steadily with worse health state as expected and there was not a serious ceiling effect, meaning that a considerable part of the patients were in the best state at start of treatment, as could be seen with health states based on the MHAQ and the EQ-5D. It might however have been a better strategy to construct the model with health states based on the EQ-5D when using the EQ-5D scores in the states. When we tested five health states based on levels from EQ-5D, 24% of patients were in the best state at start of treatment. The distribution of patients was bimodal. Our findings were in line with a previous Norwegian report concerning a ceiling effect and a bimodal shape of EQ-5D scores in patients with RA (174). In using the SF-6D for the division in health states, we could use data from ORAR and compare the initial distribution in states and distribution of utility scores for patients in NOR-DMARD and ORAR (table 12 and figures 2a-b).

Concerning the use of the SF-6D utility scores versus the EQ-5D utility scores, some assumptions have to be addressed. We assumed that a stable level in SF-6D values corresponds to a stable level in EQ-5D values. There was no EQ-5D data in ORAR. In NOR-DMARD, we had EQ-5D data for patients starting treatment with a TNFi and sDMARD. We used the results from ORAR of no change in HRQoL for patients on sDMARDs to inform the model also when using EQ-5D as outcome measure. Thus we assumed that no change in SF-6D indicated no change in EQ-5D. These MAU instruments are different. First, the range differs (0.296-1.0 in SF-6D and -0.594 to 1.0 in EQ-5D). The difference in range could result in minor changes in EQ-5D to be reduced to zero changes in SF-6D. Second, the EQ-5D is derived directly from a questionnaire with five questions of HRQoL. The SF-6D is derived from the SF-36 but disregards most of the information in the comprehensive SF-36. Only 10 of the original items in SF-36 are used in the calculation of SF-6D. It is possible that important information of HRQoL is missed in the recalculation from SF-36 to SF-6D. The disagreement between the SF-6D and the EQ-5D have been assessed in Bland-Altman plots, in relation to HAQ, and to a global health status measure (the first item of SF-36). EQ-5D was found to produce lower scores for poor health states and higher for good health states and the difference was most pronounced for patients in severe disability (HAQ ≥ 2)(174).

An argument against the use of EQ-5D is that the time trade-off (TTO) method which is the reference method to elicit values for health states in the EQ-5D, has recently been criticised (175-177). First, it is unclear whether respondents perform any meaningful TTO exercise when they respond to the EQ-5D TTO questions (175). Second, the choice of methods for developing the EQ-5D tariff have a considerable impact on the values, not least those below zero (176).

An argument in favour EQ-5D is that it is by far the most used generic HRQoL instrument worldwide, which allow for comparison with many other studies (123). A second argument in favour of EQ-5D is that the SF-6D has a floor effect with no values below 0.296 (110). Thus, we chose to base the health states on the SF-6D in order to be able to use all the available register data and we assessed utilities with both the SF-6D and the EQ-5D since the EQ-5D results are more likely to be comparable to other studies.

#### *5.2.1.2 Transition probabilities between health states*

The seven main health states, included in both strategies are based on the SF-6D, in addition to a state for death. Additionally, the TNFi-strategy includes a state for no DMARD treatment and a state for sDMARD treatment. Based on data from NOR-DMARD, patients on bDMARD therapy can stop the treatment and they might change to sDMARD therapy for different reasons. We thought it was necessary to allow these changes to occur in the model as in the real world. The assumption that patients who have stopped bDMARD therapy, reuptake a previous therapy or another bDMARD therapy one year after the stop is based on expert opinion. Patients who stop therapy in NOR-DMARD have stated (wish of) pregnancy, remission of the disease and own desire as reasons for the termination. Pregnancy is in real life a time-restricted state and we assume that patients go back to treatment after this occurrence. Patients in remission can have flares and we assume that patients in remission go back to bDMARD therapy after one year. This assumption might be wrong for some patients if they do not restart therapy after remission in one year and might give an overestimation of costs due to bDMARD use in the TNFi-strategy. We also assumed that patients who had stopped therapy because of own desire, restart therapy after one year. This assumption might overestimate the costs due to bDMARD therapy if they do not restart therapy. In the synthetic-strategy, as well as

for patients in the TNFi-strategy who have switched to sDMARD therapy, we assume that patients continue the therapy or switch to another sDMARD therapy with similar effectiveness throughout the 10-year modelling period. If some patients in real life stop sDMARD therapy, we might overestimate the costs due to sDMARD therapy. However, this overestimation will apply to both strategies.

The finding from ORAR that patients on sDMARDs did not deteriorate in utility as measured by the SF-6D over a seven-year period was somewhat surprising. Previous research has indicated that the health state of RA-patients deteriorate over time, at least in functional capacity as measured by the HAQ. For example, Wolfe and Scott *et al* both found an annual HAQ-score progression of 0.03.(137, 178) The stable level found in the ORAR data may be explained by more active use of sDMARDs, faster adjustment in therapies as well as better patient care in general, even before the introduction of bDMARDs. We assumed that this finding of no change in HRQoL over a seven-year period could be extended to 10 years, *i.e*. the time period evaluated in the NORA model. Possibly we could have assumed a deterioration from seven to 10 years, based on Wolfe's and Scott's findings (137, 178) but the reduction in HAQ-scores would then have had to be translated to HRQoL and would have introduced a new uncertainty in the estimates. Thus, our assumption of no deterioration in HRQoL over a 10-year period might have implied a bias in favour sDMARDs in the results of the cost-utility analysis.

#### *5.2.1.3 Cost estimates*

In the TNFi-strategy, we had detailed data from NOR-DMARD with up to 10 years of follow-up. A major weakness is that the data on resource use are partly self-reported as discussed above. Self-reported data may be biased by lack of recall or inaccurate data entrance. It would have been an advantage if we had had access to social security data as in the recent Swedish studies on health care costs, productivity losses and cost-effectiveness related to treatment of patients with RA (58, 159-161). More knowledge would require replicating our study with data from national electronic registers. Relevant data sources would be the Norwegian Patient Registry, the Norwegian Prescription Registry, the database of the Norwegian Health Economics Administration (HELFO) and the database of the Norwegian labour and welfare administration (NAV). In the new version of NOR-DMARD, an informed consent including the ability to link patient-level data to national registers in Norway will make such information sources available in near future.

In the synthetic-strategy we used estimates of resource use from the patients that later started with TNFi +sDMARDs, three months before the start with a bDMARD. Since the patients at this time point could have been in a particularly reduced health state, their consumption of health care at this time point could have been higher than their average consumption of health care, which might give an overestimation of the costs in the synthetic-strategy and a bias in favour bDMARDs in the results. However, we assume a constant use of health care resources over a 10-year period. Recent research indicates that hospital expenditures increased with age in Norway in the period 1998-2009 (179). If health care expenditures increase with age, the assumption of constant resource use would give an underestimation of the costs in the synthetic-strategy and a bias in favour sDMARDs. Thus, we have assumptions in the cost estimates in the synthetic-strategy that might go in opposite directions and possibly level out each other.

In one of the Swedish studies referred above, the results indicated that sick leave and disability pension increased from four years before initiation of DMARD therapy to a maximum one month after treatment start. Thereafter, the level of production loss was fairly constant for the three subsequent analysed years. Our assumption of a constant level of resource use and production loss can at least be assumed to be plausible for the first four years of our model estimates in the synthetic-strategy (160).

#### *5.2.1.4 Time perspective*

The choice of time perspective was based on two main considerations; we wanted to have a time period which was sufficiently long to allow for assessment of costs and utilities in the future, and secondly, we wanted to keep the uncertainty in the estimates at a lowest possible level. Since we had 10 years of follow-up from NOR-DMARD for costs, utilities and transition probabilities and seven years of follow-up from ORAR for utilities we chose to use a 10-year perspective. An extended time horizon would have increased the uncertainty in the estimates.

#### *5.2.1.5 Alternative strategies*

The synthetic-strategy in the NORA model includes a mixture of different sDMARD therapies, allowing for both mono- and combination therapies. Another treatment strategy that could have been considered included in the model is the specific combination of sDMARDs in triple therapy. The combination of MTX, SSZ and HCQ has recently been evaluated to have comparable effectiveness in achieving low disease activity and good functional status to biologics, in RCTs and a cohort study (162-166). The long-term effects of triple therapy compared to TNFi + MTX remain unclear.

### 5.2.2 Differences in data management in Paper II and III

The management of missing values of health care services was somewhat different in Paper II and III. At the first attempt, we assumed that the correct way of handling missing values of health care services was to perform multiple imputations using multivariate normal regression models. This is what was done for the analyses in Paper II. After further discussions with rheumatologists at the Department of Rheumatology at Diakonhjemmet hospital, we concluded that the most probable assumption to make about missing values of health care services was that no visits/stays had taken place for that particular service. Subsequently we performed imputations with zero values for missing values in the analyses in Paper III.

In the calculation of production losses, the median yearly income in Norway was used for Paper II, due to availability of data. In Paper III we used the mean income instead of the median. The mean income is recommended to use in assessments of production losses according to Norwegian guidelines (25).

## 5.3 General discussion

### 5.3.1 Paper I

The results from our study revealed that estimated cut off values for classifying a patient to be in a satisfactory health condition or as having experienced a minimal clinically important improvement vary by statistical methodology used for the assessment. The 80% specificity approach is a more conservative approach than the 75$^{th}$ percentile approach, which means that patients have to have better health and larger numerical improvements as assed by other outcome measures for being classified to be in PASS and having experienced a MCII.

We found that the cutpoints for PASS were around 0.65 in EQ-5D and 0.60 in SF-6D estimated with the 75$^{th}$ percentile approach and around 0.70 in EQ-5D and 0.65 in SF-6D estimated with the 80% specificity approach. What does such a health state correspond to? An EQ-5D value of 0.691 corresponds to a patient who has some problems walking around, no problems with self-care, some problems with performing his/her usual activities, has moderate pain or discomfort and who is not anxious or depressed. This example illustrates that patients can have problems with mobility, daily activities and moderate pain and still state that they are in a satisfactory health state. In our analysis, 30%-37%, depending on diagnosis, stated that they were in an acceptable symptom state at baseline. All of these patients had a referral from their doctor for starting or changing DMARD treatment. Intuitively, patients should not consider themselves to be in a satisfactory health state at the initiation of a DMARD-therapy. The therapy is supposed to increase the health condition in patients. It is unclear what makes a patient state that he/she is in a satisfactory health state. Many different reasons are possible, as for example positive thinking: "If I think that I am in a good health condition I will become in a good health condition in the future". It is also possible that patients state that they are in a satisfactory health state if they think they are well cared for and receive excellent treatment. Patients adapt to their disease but it has not been demonstrated that thresholds for PASS, related to disease activity change with disease duration. Thresholds for PASS were stable from initiation of DMARD treatment to a follow-up after one year in RA-patients (103).

PASS cutpoints for composite disease activity measures have been found to be in the range of moderate disease activity levels as assessed by the 75$^{th}$ percentile approach (103). A moderate disease activity level is not sufficiently low to prevent radiographic progression. Our finding of a cutpoint around 0.60-0.70 in a utility instrument indicates that the HRQoL is far from perfect health. A HRQoL of 0.70 may be a goal for a state that is acceptable and achievable for many patients with inflammatory rheumatic joint diseases but it does not reflect current ambitions with remission or low disease activity as the treatment goal (14, 180, 181).

Further, we found that the cutpoints for MCII were approximately 0 in EQ-5D and 0.02 in SF-6D with the 75$^{th}$ percentile approach and 0.18 in EQ-5D and 0.08 in SF-6D with the 80% specificity approach. The finding of no improvement in the EQ-5D and as little as 0.02 in SF-6D in patients who state that they have experienced a MCII is surprising and may indicate a weakness with the 75$^{th}$ percentile approach in identifying cutpoint values. The results indicate that a patient has experienced a MCII at zero or very small change values in utility measures. However, it can not be excluded that the utility instruments are insensitive to small health improvements. The values identified for MCII with the 80% specificity approach are more intuitively appealing, since they indicate a more substantial improvement. This is in line with results from a previous study that also compared the two methodologies for identifying MCII cutpoints in disease activity measurements. They found that the 80% specificity method performed better (106).

Finally, PASS and MCII varied with the methodology used for the assessment of cutpoint values in MAU instruments. We considered that inclusion of PASS and MCII concepts in an economic evaluation of treatment for inflammatory rheumatic joint diseases would not increase the quality, and we decided not to include them in the development of the NORA model.

### 5.3.2 Paper II

Biologic DMARD treatment entails considerable drug cost, but the total costs (including health care costs and production losses) decline during the first two years on treatment in both RA, PsA and AS. The total costs are similar across RA, PsA and AS.

The mean annual health care cost per RA patient in our study was estimated at €3,400 (sDMARD) and €19,600 (bDMARD). Our results were lower for sDMARDs and higher for bDMARDs compared to the results from a systematic review, including studies from EU and North America, which reported mean annual direct cost per patient of €4,170 (sDMARD and bDMARD patients together) (53). The recently estimated annual health care costs from Sweden were €6,352 (sDMARD and bDMARD patients together (12-28% bDMARD users depending on age group))(58). Reported annual health care costs from the US for RA patients were €9,300 in 2004 and €7,140 in 2006 (sDMARD and bDMARD) (54), thus also between our estimates for sDMARDs and bDMARDs. We chose to present costs for patients on sDMARDs and bDMARDs separately, which has not been common previously. We thought this was important since the differences in price level between sDMARD and bDMARD are large. We found that our estimated annual health care costs were lower for sDMARDs and higher for bDMARDs than the estimates from studies which report costs for sDMARDs and bDMARDs together, which is as expected.

If only the costs of the bDMARDs are considered, recent studies from the US estimated a range for the costs of TNFi from €10,853 to €17,640 per year depending on medication used (etanercept/adalimumab/infliximab) for RA patients (56, 57). Our mean yearly cost estimates for bDMARDs were €12,227; slightly over the US costs for etanercept and below the US costs for adalimumab and infliximab. Thus, it seems like our results of costs for bDMARDs are slightly lower or on the level with the cost estimates from the U.S for RA patients.

In our study, the annual health care expenditures for PsA patients on sDMARD were estimated at €3,200 and for bDMARD therapy at €18,600. When only including inpatient, outpatient, emergency department and pharmacy expenditures as in a recent study from the US (emergency department is included in inpatient stays in our study), our estimate was €16,600 for bDMARD. The corresponding costs in the U.S were estimated at €20,100 (74), thus our estimates of health care costs for PsA patients were slightly lower than in the US. If only the costs of the bDMARDs are considered, the same studies as for RA patients from the US estimated a range for the costs of TNFi from €11,396 to €20,451 per year depending on medication used (etanercept/adalimumab/infliximab) for PsA patients (56, 57). Our estimate for the bDMARDs in PsA patients was an annual mean cost of €11,957. This is in line with the US cost for etanercept and lower than for the other TNFis.

Finally, in AS patients the mean annual health care costs per patient per year was estimated at €3,700 (sDMARD) and €24,200 (bDMARD). The estimates were above the mean yearly cost estimate for AS patients from the European review, reporting a mean cost of €1,992 (sDMARD)(53). If only the costs of the bDMARDs are considered, the same studies as for RA and PsA patients from the US estimated a range for the costs of TNFi from €10,800 to €18,900 per year depending on medication used (etanercept/adalimumab/infliximab) for AS patients (56, 57). Our estimate for bDMARDs was €13,800, in line with the US results.

A comparison of costs associated with production losses are more difficult to perform since the methods used for assessing production losses varies. We assumed that full time work was the normal

workforce participation for RA patients, except for patients retired due to age and patients who had ticked the category 'other', including students, unemployed etc. This assumption might cause an overestimation of costs associated with production losses. There can be an overestimation of these costs if the normal workforce participation for a person of working age is lower than full time work (excluding students and unemployed). The estimation of a normal percentage of workforce participation is not straightforward. It can vary according to many factors, such as diagnosis, age etc.

Our estimates of mean cost of production losses per year in RA patients, assessed with the HCA, were €28,800 (sDMARD) and €41,400 (bDMARD). Similar results were found for the other diagnoses, but the production losses for AS patients on bDMARD treatment were low compared to RA and PsA. In PsA the production losses were €29,100 (sDMARD) and €37,000 (bDMARD) and in AS: €27,900 (sDMARD) and €33,500 (bDMARD). The production losses as a percentage of total costs varied between 86% (RA) and 90% (AS) for patients on sDMARDs and between 58% (AS) and 68% (RA) for patients on bDMARDs. In comparison to a European review (53), the share of production losses of total costs were higher in our study for RA patients (57% in the review) and lower for AS patients on sDMARDs (66% in the review which only included AS patients on sDMARDs).

The recent Swedish study by Eriksson et al estimated the annual costs of work loss in the prevalent RA-cohort at €16,907 (sDMARD and bDMARD) applying the HCA (58). The Swedish estimate is approximately 60% of our estimate for patients on sDMARDs and 40% of our estimate for patients on bDMARDs. Two main differences in the cost assessments might contribute to the differences in results. First, we applied an estimate of the social costs of labour of 40% while the Swedish application of social fees was 31%. Further, the annual income estimate in Norway was €50,245 and in Sweden, the average monthly salary was €2,990, yielding an annual estimate of €35,880. The Swedish estimate included the social fees. Thus, the estimated income in Norway is substantially higher than the estimated income in Sweden. A similarity between the studies is the share of production loss of the total costs incurred by RA-patients. In our study, production loss constituted 86% (sDMARDs) and 68% (bDMARDs) of the total costs and in the Swedish study, production loss was 73% (sDMARD and bDMARD) of total costs (58).

In conclusion, both health care costs and production losses in patients with inflammatory rheumatic joint diseases seem to be high in Norway and are probably higher in Norway than in other European countries. Previous studies have reported lower cost estimates than our study, except from recent studies performed in the US This is probably due to a high cost level overall in Norway, to our study being performed more recently than most comparative studies and finally that we include the costs of bDMARDs which is not the case in all previous studies.

### 5.3.3 Paper III

TNFi-treatment for RA is cost-effective when accounting for production losses, assuming a Norwegian willingness-to-pay level of approximately €67,300. Excluding production losses, TNFi-treatment is cost-effective using EQ-5D, but not SF-6D. In the following, I will discuss the findings from our model estimates compared to results from previous studies. Further, I will discuss differences between our model and previous models. Finally, the introduction of biosimilar bDMARDs and their consequences for the estimated ICER is discussed.

Previous studies have varied quite substantially in ICERs for TNFi versus sDMARDs in patients with RA. Our results were in the lower part of the reported range of ICERs for bDMARDs versus sDMARDs in RA patients who have failed ≥2DMARDs, in a recently published systematic review (144).

Our estimate for the incremental cost per QALY including health care costs and EQ-5D utilities was €61,285, close to an estimate from the UK for two years of infliximab therapy of €56,100 (131). This study by Kobelt and co-workers used unit costs from 1999 and 2000 for the UK version and our unit costs were from 2012. The same study reported results of ICERs also based on Swedish prices (from 2001 and 2002) and these were lower than our results (€44,500) which is not unexpected with unit costs from 10 years before our study (131).

A later study by Kobelt and co-workers (136) presented results that were very similar to our estimates using EQ-5D utilities and including production losses. The incremental costs over 10 years when a patient used etanercept + MTX instead of MTX alone was €42,148 (€38,641 in our study) and the incremental QALY gain was 0.91 (0.97 in our study). The ICER in the study by Kobelt and co-workers was €46,494 in a 10-year perspective and in our study the corresponding ICER was €39,841 (136). Reasons for the slightly higher ICER in the previous study might be that Kobelt and co-workers included informal help and care as well as loss of leisure time in the cost estimates, which we did not. Another difference was that we had longitudinal data on treatment effectiveness up to 10 years while they used trial based data with two years of follow-up and extrapolated transition probabilities in the model based on the second year of the trial.

A model from the Netherlands reported an ICER for TNFi compared to usual treatment for RA patients (132). Usual treatment at the time of the study was sulfasalazine followed by MTX. The reported ICER of €163,556 (estimated with EQ-5D utilities and including production losses) was much higher than our ICER of €39,841 (EQ-5D utilities and production losses).

Compared to a study based on Swedish data, evaluating the effectiveness of adalimumab with HUI-3, the ICERs from our study based on the EQ-5D and applying health care cost, were included in the range of sensitivity analysis results from the adalimumab study. Our estimates applying SF-6D were higher than reported in this study (139). A more recent Swedish study reported a base case ICER of €22,830 for infliximab versus sDMARD therapy including production losses and using EQ-5D utility estimates from the HAQ and DAS28. In the base case, an annual HAQ-score progression of 0.065 was assumed. However when a lower progression rate was assumed, *i.e.* 0.031, the ICER rose to just over €35,000. Our result of €39,841 is close to €35,000. The results are thus quite similar if a lower HAQ-score progression rate than 0.065 is assumed in the Markov model reported by Lekander and co-workers (142).

In previous studies, direct application of data from MAU instruments is rare. One study reported using HUI-3 data from RCTs for the effectiveness of adalimumab, but for the effectiveness of the comparators (infliximab and etanercept) a transformation was made from the HAQ to the HUI-3 (139). Another study used SF-6D data directly generated from SF-36 but for the estimation of EQ-5D utilities a mapping was made from the HAQ (141). Most other studies have used a transformation from the HAQ to generate utilities (131, 135, 136, 138, 142). In our study we had direct utility estimates from the EQ-5D and from the SF-36 to the SF-6D in NOR-DMARD. In ORAR we had utility estimates from the SF-36 to SF-6D.

Similarly, long time follow-up data is rare. Only one study reported the same length of follow up as we had, *i.e*. 10 years (142).

In the NORA model, we combine direct data from MAU instruments with 10 years of follow-up. This combination is unique in cost-utility analyses of TNFi treatment in RA. Further, we introduce a new way of defining health states by dividing health states into levels of HRQoL. Thus, data do not need to be transformed from scores of function or disease activity to QALYs.

Assuming a prevalence of 0.44% of RA and a population of 5, 096,300, there are approximately 22,424 RA-patients in Norway. Of these, 59.5% use DMARDs (151). Multiplying the number of RA-patients using DMARDs (13,342) with the EVPI (€201) (NOK 1,494) per patient, gives a total value of €2,681,765 (NOK 19,933,117). This is the estimation of the value of performing additional research on the 78 parameters with distributions in the model. Assuming a WTP threshold of €67,300 (NOK 500,000) the estimated sum corresponds to an opportunity cost of approximately 40 QALYs in the base case analysis. It is unlikely that it would be possible to gain perfect information in the uncertain parameter values and even more unlikely at less than NOK 20,000,000.

Biosimilar bDMARDs represent a new opportunity for reduction of drugs costs. The reduced costs of biosimilars represent a particularly important opportunity in countries with poor financial status. It has been shown that access to and restrictions in the prescription of bDMARDs are related to gross domestic product (182, 183). Biosimilars are not exact copies of the innovator products but have undergone a stringent regulatory process before approval (184, 185). Biosimilar infliximab from Celltrion has been tested in two large trials, PLANETRA and PLANETAS (186, 187) and the European Medicines Agency (EMA) and the Norwegian regulatory agency approved in 2013 the drug for these indications and approved extrapolation to PsA, psoriasis and inflammatory bowel diseases. Orion Pharma offered in the last LIS tender for 2014 a discount of 39% for Remsima® compared to the innovator infliximab (Remicade®). The cost for Remsima® for the first year in a patient of 75 kg and with standard dose of 3 mg/kg will be between €6,000 and €7,000 (NOK 51,588). Thus, Remsima® is the recommended option when starting a new treatment. However, substitution of Remicade® with Remsima® in patients on stable treatment with the innovator drugs is not recommended until more data is available. The Norwegian government has in the budget for 2014 reserved NOK 20 mill (about €2,5 mill) for a study to examine switching from the innovator to biosimilar infliximab.

In our study, we used prices from 2012 for all cost items, including drugs. The price of Remsima® in 2014 is 45% lower than the price of Remicade® in 2012 for the first year in a patient with RA of 75 kg and with standard dose of 3 mg/kg. If all patients on TNFis in the model use biosimilars or innovator drugs to the same reduction in price (-45%) and the prices on sDMARDs and other cost items are assumed stable, an approximate ICER for the base case including health care costs and EQ-5D utilities is €30,000 (NOK 240,000). This approximation includes an estimation of the part of DMARDs amounting to 53% of the health care costs, which is the part of DMARD cost for RA-patients in the average health state at start of treatment with TNFi.

# 6 POLICY IMPLICATIONS

Paper I: The concepts of PASS and MCII need to be better founded in research to be used in effectiveness evaluations.

Paper II: The wide cost spectrum found when examining costs related to RA, PsA and AS indicates that it is important to consider a range of possible cost sources when performing cost assessments and budget impact analyses in inflammatory rheumatic joint diseases.

Paper III: Paper III addresses the issue of whether society should reimburse treatment with bDMARDs for RA. In Norway, current priority recommendations use three criteria: severity of the health state, effectiveness of treatment and finally cost-effectiveness. The combination of the three criteria provides the basis for prioritizing. In the NORA model, we have assessed one of the three criteria and found that the cost-effectiveness of TNFi-treatment versus sDMARD treatment for RA-patients based on clinical real life data in Norway seems to be within the range of previously funded treatments. RA is a severe disease, the effectiveness of treatment with TNFis is well documented and our results indicate that bDMARDs are cost-effective. In conclusion, the results from this study indicate that TNFi-treatment should continue to be reimbursed for patients with RA in Norway.

We found that the choice of MAU instrument influences the results considerable. Since both the EQ-5D and the SF-6D have weaknesses and no consensus exist of which instrument to use it is important to present results with both. In future cost-effectiveness analyses, more than one MAU instrument should be used when possible, until consensus in which instrument to prefer has been reached.

# 7 FUTURE RESEARCH

All three papers in this thesis raise issues that warrant further research. It would be needed to do qualitative studies and explore patients' considerations about PASS and MCII. Also, determinants of PASS and MCII would need further studies. One hypothesis would be that whether patients are satisfied with their current health status and treatment improvement depends on patient characteristics such as disease duration as well as the status of the disease.

Paper II was mainly based on questionnaire data from patients or health personnel. Such data may be biased by lack of recall or inaccurate data entrance. More knowledge would require replicating the cost study with data from national electronic registers. Relevant data sources would be the Norwegian Patient Registry, the Norwegian Prescription Registry, the database of the Norwegian Health Economics Administration (HELFO) and the database of the Norwegian labour and welfare administration (NAV). In the new version of NOR-DMARD, an informed consent including the ability to link patient-level data to national registers in Norway will make such information sources available in near future.

The NORA model can be further developed in different directions. First, the model could be expanded to cover cost-utility analyses for patients with PsA and AS. Since the models' health states are based on HRQoL, a use of the model for other diagnoses in which HRQoL data is collected is possible. Second, in RA patients it is relevant to limit the analysis to patient subgroups, such as only including patients who have failed one sDMARD before the initiation of bDMARD in the intervention group. The comparator could be continued use of sDMARDs or patients who had failed two sDMARDs before the introduction of bDMARDs. Using propensity-score analysis could be a way of reducing initial differences between comparative patient groups. Third, if data become available, triple therapy (MTX, SSZ and HCQ) is a relevant treatment arm to include. Fourth, the model could be expanded to explore differences between different bDMARDs to the extent there is evidence that they are clinically different. Fifth, the model could be revised to capture specific treatment switches

between different bDMARDs. This would likely require a development from a Markov model to a discrete event simulation model since the number of states would increase extensively (32). Sixth, if data from RCTs become available, the data would be very relevant to include in the model, which would further validate the results.

# 8 REFERENCES


1. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 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.
2. Scott DL, Coulton BL, Symmons DPM, Popert AJ. Long-term outcome of treating rheumatoid arthritis: results after 20 years. The Lancet. 1987;329(8542):1108-11.
3. Mitchell DM, Spitz PW, Young DY, Bloch DA, McShane DJ, Fries JF. Survival, prognosis, and causes of death in rheumatoid arthritis. Arthritis Rheum. 1986;29(6):706-14.
4. Scott DL, Wolfe F, Huizinga TWJ. Rheumatoid arthritis. The Lancet. 2010;376(9746):1094-108.
5. Amherd-Hoekstra A, Naher H, Lorenz HM, Enk AH. Psoriatic arthritis: a review. J Dtsch Dermatol Ges. 2010;8(5):332-9.
6. Braun J, Sieper J. Ankylosing spondylitis. Lancet. 2007;369(9570):1379-90.
7. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. The Lancet. 2004;363(9410):675-81.
8. Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, et al. Infliximab (chimeric anti-tumour necrosis factor@monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. The Lancet. 1999;354(9194):1932-9.
9. Breedveld F, Weisman M, Kavanaugh A, Cohen S, 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.
10. Knevel R, Schoels M, Huizinga TWJ, Aletaha D, Burmester GR, Combe B, et al. Current evidence for a strategic approach to the management of rheumatoid arthritis with disease-modifying antirheumatic drugs: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis. Ann Rheum Dis. 2010;69(6):987-94.
11. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. Cochrane Database Syst Rev. 2009(4):Cd007848.
12. Sosial- og helsedepartementet [Ministry of Health and Care Services]. Prioritering på ny. Gjennomgang av retningslinjer for prioriteringer innen norsk helsetjeneste. [Priorities again. Review of guidelines for priorities in Norwegian health care.]. Departementenes servicesenter; 2012; NOU 1997:18.
13. Sosial- og helsedepartementet [Ministry of Health and Care Services]. Retningslinjer for prioriteringer innen norsk helsetjeneste. [Guidelines for priorities in Norwegian health care.]. Departementenes servicesenter; 1987; NOU 1987:23.
14. 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.
15. Samuelson P, Nordhaus W. Economics. Twelfth ed. New York: McGraw-Hill; 1985.
16. Stieglitz J, Walsh CE. Principles of microeconomics. Third ed. New York: W.W. Norton & Company, Inc.; 2002.
17. Folland S, Goodman AC, Stano M. The Economics of Health and Health Care. 4th ed. New Jersey: Pearson Prentice Hall; 2004.
18. Drummond MF, Sculpher MJ, Stoddart GL, O'Brien BJ, Torrance GW. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford: Oxford University Press; 2005.
19. Brouwer WB, Culyer AJ, van Exel NJ, Rutten FF. Welfarism vs. extra-welfarism. J Health Econ. 2008;27(2):325-38.

20. Hurley J. An overview of the normative economics of the health care sector. In: Culyer AJ, Newhouse JP, editors. Handbook of Health Economics. Amsterdam: Elsevier Science; 2000. p. 55-118.
21. Hurley J. Welfarism, extra-welfarism and evaluative economic anaysis in the health care sector. Barer M, Getzen T, Stoddard G, editors. Chichester: John Wiley & Sons Ltd.; 1998.
22. Johansson P-O. An introduction to modern welfare economics. Cambridge: Cambridge University Press; 1991.
23. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press; 1996.
24. Hunink M, Glasziou P, Siegel J, Weeks J, Pliskin J, Elstein A, et al. Decision making in health and medicine: intergrating evidence and values. Cambridge: Cambridge University Press; 2001.
25. Helsedirektoratet [The Norwegian Directorate of Health]. Økonomisk evaluering av helsetiltak – en veileder [Economic evaluation of healthcare interventions - a guide]. Oslo: 2012.
26. Olsen JA, Richardson J. Production gains from health care: what should be included in cost-effectiveness analyses? Soc Sci Med. 1999;49(1):17-26.
27. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. J Health Econ. 1995;14(2):171-89.
28. Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ. Discounting and decision making in the economic evaluation of health-care technologies. Health Econ. 2011;20(1):2-15.
29. Briggs A, Claxton K, Schulper M. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press; 2006.
30. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model Parameter Estimation and Uncertainty Analysis: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. Med Decis Making. 2012;32(5):722-32.
31. Stinnett AA, Paltiel AD. Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means. Med Decis Making. 1997;17(4):483-9.
32. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Moller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--4. Value Health. 2012;15(6):821-7.
33. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–3. Med Decis Making. 2012;32(5):690-700.
34. Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. The Lancet. 2009;373(9664):659-72.
35. Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: Identifying reversible and irreversible components. Arthritis Rheum. 2006;54(9):2784-93.
36. Smolen JS, Aletaha D. Patients with rheumatoid arthritis in clinical care. Ann Rheum Dis. 2004;63(3):221-5.
37. Lillegraven S, van der Heijde D, Uhlig T, Kvien TK, Haavardsholm EA. What is the clinical relevance of erosions and joint space narrowing in RA? Nature reviews Rheumatology. 2012;8(2):117-20.
38. 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.
39. Nikolaisen C, Besada E, Nossent J. Nye kriterier for revmatoid artritt. Tidsskr Nor Legeforen. 2012;132(2):175-7.
40. 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.
41. 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.
42. Alamanos Y, Voulgari PV, Drosos AA. Incidence and Prevalence of Rheumatoid Arthritis, Based on the 1987 American College of Rheumatology Criteria: A Systematic Review. Semin Arthritis Rheum. 2006;36(3):182-8.

43. Uhlig T, Kvien TK. Is rheumatoid arthritis disappearing? Ann Rheum Dis. 2005;64(1):7-10.
44. Kvien TK, Glennas A, Knudsrod OG, Smedstad LM, Mowinckel P, Forre O. The prevalence and severity of rheumatoid arthritis in Oslo. Results from a county register and population survey. Scand J Rheumatol. 1997;26:412-8.
45. Smolen JS, Landewé 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;69:964-75.
46. Gibofsky A, Yazici Y. Treatment of rheumatoid arthritis: strategies for achieving optimal outcomes. Ann Rheum Dis. 2010;69(6):941-2.
47. Roussy J-P, Bessette L, Rahme E, Bernatsky S, Légaré J, Lachaine J. Rheumatoid arthritis pharmacotherapy and predictors of disease-modifying anti-rheumatic drug initiation in the early years of biologic use in Quebec, Canada. Rheumatol Int. 2014;34(1):75-83.
48. Patient.co.uk. [Cited 2014 May 30]. Available from: http://www.patient.co.uk/doctor/disease-modifying-antirheumatic-drugs-dmards-pro.
49. Nandi P, Kingsley GH, Scott DL. Disease-modifying antirheumatic drugs other than methotrexate in rheumatoid arthritis and seronegative arthritis. Curr Opin Rheumatol. 2008;20(3):251-6.
50. U.S. Food and Drug Administration. What is a biological product? [updated 2010 Nov 05; cited 2013 Nov 23]. Available from: http://www.fda.gov/BiologicsBloodVaccines/ResourcesforYou/default.htm.
51. Smolen J, Aletaha D, Koeller M, Weisman M, Emery P. New therapies for treatment of rheumatoid arthritis. Lancet. 2007;370(9602):1861-74.
52. Fautrel B, Verstappen SM, Boonen A. Economic consequences and potential benefits. BestPractResClinRheumatol. 2011;25(4):607-24.
53. Franke LC, Ament AJ, van de Laar MA, Boonen A, Severens JL. Cost-of-illness of rheumatoid arthritis and ankylosing spondylitis. ClinExpRheumatol. 2009;27(4 Suppl 55):S118-S23.
54. Simons WR, Rosenblatt LC, Trivedi DN. The economic consequences of rheumatoid arthritis: analysis of Medical Expenditure Panel Survey 2004, 2005, and 2006 data. J Occup Environ Med. 2012;54(1):48-55.
55. Lee T-J, Park BH, Son HK, Song R, Shin KC, Lee EB, et al. Cost of Illness and Quality of Life of Patients with Rheumatoid Arthritis in South Korea. Value Health. 2012;15(1, Supplement):S43-S9.
56. Schabert VF, Watson C, Gandra SR, Goodman S, Fox KM, Harrison DJ. Annual costs of tumor necrosis factor inhibitors using real-world data in a commercially insured population in the United States. J Med Econ. 2012;15(2):264-75.
57. Bonafede MK, Gandra S, Watson C, Princic N, Fox K. Cost per Treated Patient for Etanercept, Adalimumab, and Infliximab Across Adult Indications: a Claims Analysis. Adv Ther. 2012;29(3):234-48.
58. Eriksson JK, Johansson K, Askling J, Neovius M. Costs for hospital care, drugs and lost work days in incident and prevalent rheumatoid arthritis: how large, and how are they distributed? Ann Rheum Dis. 2013 Dec 9; doi: 10.1136/annrheumdis-2013-204080. [Epub ahed of print].
59. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis. 2009;68(6):777-83.
60. Bakland G, Nossent HC. Epidemiology of spondyloarthritis: a review. Curr Rheumatol Rep. 2013;15(9):351.
61. Eder L, Gladman DD. Psoriatic arthritis: phenotypic variance and nosology. Curr Rheumatol Rep. 2013;15(3):316.
62. Taylor W, Gladman D, Helliwell P, Marchesoni A, Mease P, Mielants H. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum. 2006;54(8):2665-73.
63. Alamanos Y, Voulgari PV, Drosos AA. Incidence and prevalence of psoriatic arthritis: a systematic review. J Rheumatol. 2008;35(7):1354-8.

64. Hoff M, Gulati AM, Romundstad PR, Kavanaugh A, Haugeberg G. Prevalence and incidence rates of psoriatic arthritis in central Norway: data from the Nord-Trondelag Health Study (HUNT). Ann Rheum Dis. 2013 Aug 20; doi: 10.1136/annrheumdis-2013-203862. [Epub ahead of print].
65. Gladman DD. Psoriatic arthritis. Dermatol Ther. 2004;17(5):350-63.
66. Gossec L, Smolen JS, Gaujoux-Viala C, Ash Z, Marzo-Ortega H, van der Heijde D, et al. European League Against Rheumatism recommendations for the management of psoriatic arthritis with pharmacological therapies. Ann Rheum Dis. 2012;71(1):4-12.
67. Ash Z, Gaujoux-Viala C, Gossec L, Hensor EM, FitzGerald O, Winthrop K, et al. A systematic literature review of drug therapies for the treatment of psoriatic arthritis: current evidence and meta-analysis informing the EULAR recommendations for the management of psoriatic arthritis. Ann Rheum Dis. 2012;71(3):319-26.
68. Mease PJ, Goffe BS, Metz J, VanderStoep A, Finck B, Burge DJ. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. The Lancet. 2000;356(9227):385-90.
69. Mease PJ, Gladman DD, Ritchlin CT, Ruderman EM, Steinfeld SD, Choy EHS, et al. Adalimumab for the treatment of patients with moderately to severely active psoriatic arthritis: Results of a double-blind, randomized, placebo-controlled trial. Arthritis Rheum. 2005;52(10):3279-89.
70. Kavanaugh A, McInnes I, Mease P, Krueger GG, Gladman D, Gomez-Reino J, et al. Golimumab, a new human tumor necrosis factor α antibody, administered every four weeks as a subcutaneous injection in psoriatic arthritis: Twenty-four–week efficacy and safety results of a randomized, placebo-controlled study. Arthritis Rheum. 2009;60(4):976-86.
71. Antoni CE, Kavanaugh A, Kirkham B, Tutuncu Z, Burmester GR, Schneider U, et al. Sustained benefits of infliximab therapy for dermatologic and articular manifestations of psoriatic arthritis: Results from the infliximab multinational psoriatic arthritis controlled trial (IMPACT). Arthritis Rheum. 2005;52(4):1227-36.
72. McInnes IB, Kavanaugh A, Gottlieb AB, Puig L, Rahman P, Ritchlin C, et al. Efficacy and safety of ustekinumab in patients with active psoriatic arthritis: 1 year results of the phase 3, multicentre, double-blind, placebo-controlled PSUMMIT 1 trial. Lancet. 2013;382(9894):780-9.
73. Ritchlin C, Rahman P, Kavanaugh A, McInnes IB, Puig L, Li S, et al. Efficacy and safety of the anti-IL-12/23 p40 monoclonal antibody, ustekinumab, in patients with active psoriatic arthritis despite conventional non-biological and biological anti-tumour necrosis factor therapy: 6-month and 1-year results of the phase 3, multicentre, double-blind, placebo-controlled, randomised PSUMMIT 2 trial. Ann Rheum Dis. 2014 Jan 30; doi: 10.1136/annrheumdis-2013-204655. [Epub ahead of print].
74. Zhu B, Edson-Heredia E, Gatz JL, Guo J, Shuler CL. Treatment patterns and health care costs for patients with psoriatic arthritis on biologic therapy: a retrospective cohort study. Clin Ther. 2013;35(9):1376-85.
75. van Tubergen A, van der Heijde D, Dougados M, Mielants H, Landewé R. Are syndesmophytes most prevalent in the lumbar or in the cervical spine in patients with ankylosing spondylitis and do they develop in a specific direction? Rheumatology (Oxford). 2012;51(8):1432-9.
76. Wanders A, Landewe R, Dougados M, Mielants H, van der Linden S, van der Heijde D. Association between radiographic damage of the spine and spinal mobility for individual patients with ankylosing spondylitis: can assessment of spinal mobility be a proxy for radiographic evaluation? Ann Rheum Dis. 2005;64(7):988-94.
77. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum. 1984;27(4):361-8.
78. Dougados M, van der Linden S, Juhlin R, Huitfeldt B, Amor B, Calin A, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. Arthritis Rheum. 1991;34(10):1218-27.
79. Amor B, Dougados M, Listrat V, Menkes CJ, Roux H, Benhamou C, et al. Are classification criteria for spondylarthropathy useful as diagnostic criteria? Rev Rhum Engl Ed. 1995;62(1):10-5.
80. Bakland G, Nossent HC, Gran JT. Incidence and prevalence of ankylosing spondylitis in Northern Norway. Arthritis Rheum. 2005;53(6):850-5.

81. Hukuda S, Minami M, Saito T, Mitsui H, Matsui N, Komatsubara Y, et al. Spondyloarthropathies in Japan: nationwide questionnaire survey performed by the Japan Ankylosing Spondylitis Society. J Rheumatol. 2001;28(3):554-9.
82. Haglund E, Bremander AB, Petersson IF, Strombeck B, Bergman S, Jacobsson LT, et al. Prevalence of spondyloarthritis and its subtypes in southern Sweden. Ann Rheum Dis. 2011;70(6):943-8.
83. Feldtkeller E, Khan MA, van der Heijde D, van der Linden S, Braun J. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. Rheumatol Int. 2003;23(2):61-6.
84. Dean LE, Jones GT, Macdonald AG, Downham C, Sturrock RD, Macfarlane GJ. Global prevalence of ankylosing spondylitis. Rheumatology (Oxford). 2013 Dec 9; doi: 10.1093/rheumatology/ket387. [Epub ahead of print].
85. Braun J, Bollow M, Remlinger G, Eggens U, Rudwaleit M, Distler A, et al. Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. Arthritis Rheum. 1998;41(1):58-67.
86. Zochling J, van der Heijde D, Burgos-Vargas R, Collantes E, Davis JC, Jr., Dijkmans B, et al. ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis. 2006;65(4):442-52.
87. Braun J, van den Berg R, Baraliakos X, Boehm H, Burgos-Vargas R, Collantes-Estevez E, et al. 2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. Ann Rheum Dis. 2011;70(6):896-904.
88. Boers M, Brooks P, Strand C, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology. J Rheumatol. 1998;25(2):198-9.
89. Aletaha D, Landewe R, Karonitsch T, Bathon J, Boers M, Bombardier C, et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Ann Rheum Dis. 2008;67(10):1360-4.
90. Aletaha D, Ward MM, Machold KP, Nell VPK, Stamm T, Smolen JS. Remission and active disease in rheumatoid arthritis: Defining criteria for disease activity states. Arthritis Rheum. 2005;52(9):2625-36.
91. Fries JF, Spitz P, Kraines G, Holman R. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137-45.
92. Pincus T, Summey J, Soraci JR S, Wallstone K, Hummon N. Assessment of patient satisfaction in activities of daily living using a modified stanford health assessment questionnaire. Arthritis Rheum. 1983;26(11):1346-53.
93. 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.
94. van der Heijde DM, van't Hof MA, van Riel PL, van Leeuwen MA, van Rijswijk MH, van de Putte LB. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. Ann Rheum Dis. 1992;51(2):177-81.
95. Prevoo MLL, Van'T Hof MA, Kuper HH, Van Leeuwen MA, Van De Putte LBA, Van Riel PLCM. Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38(1):44-8.
96. Walter MJ, Mohd Din SH, Hazes JM, Lesaffre E, Barendregt PJ, Luime JJ. Is Tightly Controlled Disease Activity Possible with Online Patient-reported Outcomes? J Rheumatol. 2014 Feb 5; doi: 10.3899/jrheum.130174. [Ebup ahead of print].
97. van Riel PL, van Gestel AM. Clinical outcome measures in rheumatoid arthritis. Ann Rheum Dis. 2000;59 Suppl 1:i28-31.
98. Lie E. Effectiveness of synthetic and biological disease modifying antirheumatic drugs in patients with inflammatory joint diseases. Results from the NOR-DMARD register [doctoral thesis]. Oslo: University of Oslo, Faculty of Medicine; 2012.

99. Sørensen J, Linde L, Ostergaard M, Hetland ML. Quality-Adjusted Life Expectancies in Patients with Rheumatoid Arthritis - Comparison of Index Scores from EQ-5D, 15D, and SF-6D. Value Health. 2012;15(2):334-9.
100. Uhlig T, Haavardsholm EA, Kvien TK. Comparison of the Health Assessment Questionnaire (HAQ) and the modified HAQ (MHAQ) in patients with rheumatoid arthritis. Rheumatology (Oxford). 2006;45(4):454-8.
101. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis. 2007;66(Suppl 3):40-1.
102. Tubach F, Ravaud P, Beaton D, Boers M, Bomardier C, Felson DT, et al. Minimal Clinically Important Improvement and Patient Acceptable Symptom State for Subjective Outcome Measures in Rheumatic Disorders. J Rheumatol. 2007;34(5):1188-93.
103. Heiberg T, Kvien TK, Mowinckel P, Aletaha D, Smolen JS, Hagen KB. Identification of disease activity and health status cut-off points for the symptom state acceptable to patients with rheumatoid arthritis. Ann Rheum Dis. 2008;67(7):967-71.
104. Maksymowych WP, Richardson R, Mallon C, van der Heijde D, Boonen A. Evaluation and validation of the patient acceptable symptom state (PASS) in patients with ankylosing spondylitis. Arthritis Rheum. 2007;57(1):133-40.
105. Tubach F, Pham T, Skomsvoll JF, Mikkelsen K, Bjørneboe O, Ravaud P, et al. Stability of the patient acceptable symptomatic state over time in outcome criteria in ankylosing spondylitis. Arthritis Rheum. 2006;55(6):960-4.
106. Aletaha D, Funovits J, Ward MM, Smolen JS, Kvien TK. Perception of improvement in patients with rheumatoid arthritis varies with disease activity levels at baseline. Arthritis Rheum. 2009;61(3):313-20.
107. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. J Clin Epidemiol. 2002;55(9):900-8.
108. Kvien TK, Heiberg MS, Lie E, Kaufmann C, Mikkelsen K, Nordvåg B-Y, et al. A Norwegian DMARD register: Prescriptions og DMARDs and biological agents to patients with inflammatory rheumatic diseases. Clin Exp Rheumatol. 2005;23(Suppl. 39):s188-94.
109. Ware JE, Sherbourne CD. The MOS 36-ltem Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. Med Care. 1992;30(6):473-83.
110. 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.
111. Dolan P. Modeling Valuations for EuroQol Health States. Med Care. 1997;35(11):1095-108.
112. Brazier J, Ratcliffe J, Salomon J, A, Tsuchiya A. Measuring and Valuing Health Benefits for Economic Evaluation. New York: Oxford University Press, Inc.; 2007.
113. Kosinski M, Keller SD, Hatoum HT, Sheldon XK, Ware JE, Jr. The SF-36 Health Survey as a Generic Outcome Measure in Clinical Trials of Patients with Osteoarthritis and Rheumatoid Arthritis: Tests of Data Quality, Scaling Assumptions and Score Reliability. Med Care. 1999;37(5):MS10-MS22.
114. Ruta DA, Hurst NP, Kind P, Hunter M, Stubbings A. Measuring health status in British patients with rheumatoid arthritis: reliability, validity and responsiveness of the short form 36-item health survey (SF-36). Rheumatology (Oxford). 1998;37(4):425-36.
115. Kosinski M, Keller SD, Ware JE, Jr., Hatoum HT, Sheldon XK. The SF-36 Health Survey as a Generic Outcome Measure in Clinical Trials of Patients with Osteoarthritis and Rheumatoid Arthritis: Relative Validity of Scales in Relation to Clinical Measures of Arthritis Severity. Med Care. 1999;37(5):MS23-MS39.
116. Loge JH, Kaasa S, Hjermstad MJ, Kvien TK. Translation and Performance of the Norwegian SF-36 Health Survey in Patients with Rheumatoid Arthritis. I. Data Quality, Scaling Assumptions, Reliability, and Construct Validity. J Clin Epidemiol. 1998;51(11):1069-76.
117. Wells G, Li T, Maxwell L, Maclean R, Tugwell P. Responsiveness of patient reported outcomes including fatigue, sleep quality, activity limitation, and quality of life following treatment with abatacept for rheumatoid arthritis. Ann Rheum Dis. 2008;67(2):260-5.

118. Linde L, Sorensen J, Ostergaard M, Horslev-Petersen K, Hetland ML. Health-related quality of life: validity, reliability, and responsiveness of SF-36, 15D, EQ-5D [corrected] RAQoL, and HAQ in patients with rheumatoid arthritis. J Rheumatol. 2008;35(8):1528-37.
119. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ. 2004;13(9):873-85.
120. Marra CA, Rashidi AA, Guh D, Kopec JA, Abrahamowicz M, Esdaile JM, et al. Are indirect utility measures reliable and responsive in rheumatoid arthritis patients? Qual Life Res. 2005;14(5):1333-44.
121. Russell AS, Conner-Spady B, Mintz B, Mallon C, W.P M. The Responsiveness of Generic Health Status Measures As Assessed in Patients with RA Receiving Infliximab. J Rheumatol. 2003;30(5):941-7.
122. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.
123. Richardson J. A comparison of 6 multi attribute utility instruments in 7 disease areas. Population Health Strategic Reseach, Seminar Series. Deakin: Deakin Univeristy; 2013. .
124. Brooks R. EuroQol: the current state of play. Health Policy. 1996;37(1):53-72.
125. Herdman M, Gudex C, Lloyd A, Janssen MF, 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.
126. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Rheumatology (Oxford). 1997;36(5):551-9.
127. Helse- og omsorgsdepartementet [the Ministry of Health and Care Services]. Proposisjon til Stortinget for budsjettåret 2014 [Proposition to the Parliament for the fiscal year 2014]. 2013.
128. Kavanaugh A, Heudebert G, Cush J, Jain R. Cost evaluation of novel therapeutics in rheumatoid arthritis (CENTRA):A decision analysis model. Semin Arthritis Rheum. 1996;25(5):297-307.
129. Wong JB, Singh G, Kavanaugh A. Estimating the cost-effectiveness of 54 weeks of infliximab for rheumatoid arthritis. The American Journal of Medicine. 2002;113(5):400-8.
130. Wong JB, Ramey DR, Singh G. Long-term morbidity, mortality, and economics of rheumatoid arthritis. Arthritis Rheum. 2001;44(12):2746-9.
131. Kobelt G, Jönsson L, Young A, Eberhardt K. The cost–effectiveness of infliximab (Remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study. Rheumatology (Oxford). 2003;42(2):326-35.
132. Welsing PMJ, Severens JL, Hartman M, van Riel PLCM, Laan RFJM. Modeling the 5-year cost effectiveness of treatment strategies including tumor necrosis factor-blocking agents and leflunomide for treating rheumatoid arthritis in the Netherlands. Arthritis Rheum. 2004;51(6):964-73.
133. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumar necrosis factor in rheumatoid arthritis. Health Technol Assess. 2004;8(11):1-91.
134. Jobanputra P, Barton P, Bryan S, Burls A. The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation. Health Technol Assess. 2002;6(21):1-110.
135. Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. Rheumatology (Oxford). 2004;43(1):62-72.
136. Kobelt G, Lindgren P, Singh A, Klareskog L. Cost effectiveness of etanercept (Enbrel) in combination with methotrexate in the treatment of active rheumatoid arthritis based on the TEMPO trial. Ann Rheum Dis. 2005;64(8):1174-9.
137. Scott DL, Pugner K, Kaarela K, Doyle DV, Woolf A, Holmes J, et al. The links between joint damage and disability in rheumatoid arthritis. Rheumatology (Oxford). 2000;39(2):122-32.
138. Barbieri M, Wong JB, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. Pharmacoeconomics. 2005;23(6):607-18.

139. Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. Ann Rheum Dis. 2005;64(7):995-1002.
140. Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. Health Technol Assess. 2006;10(42):iii-iv, xi.
141. Brennan A, Bansback N, Nixon R, Madan J, Harrison M, Watson K, et al. Modelling the cost effectiveness of TNF-α antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. Rheumatology (Oxford). 2007;46(8):1345-54.
142. Lekander I, Borgström F, Svarvar P, Ljung T, Carli C, van Vollenhoven RF. Cost-effectiveness of real-world infliximab use in patients with rheumatoid arthritis in Sweden. Int J Technol Assess Health Care. 2010;26(01):54-61.
143. Kobelt G, Lindgren P, Lindroth Y, Jacobson L, Eberhardt K. Modelling the effect of function and disease activity on costs and quality of life in rheumatoid arthritis. Rheumatology (Oxford). 2005;44(9):1169-75.
144. van der Velde G, Pham B, Machado M, Ieraci L, Witteman W, Bombardier C, et al. Cost-effectiveness of biologic response modifiers compared to disease-modifying antirheumatic drugs for rheumatoid arthritis: A systematic review. Arthritis Care Res. 2011;63(1):65-78.
145. Modena V, Bianchi G, Roccatello D. Cost-effectiveness of biologic treatment for rheumatoid arthritis in clinical practice: An achievable target? Autoimmunity Reviews. 2013;12(8):835-8.
146. Tosh J, Brennan A, Wailoo A, Bansback N. The Sheffield rheumatoid arthritis health economic model. Rheumatology (Oxford). 2011;50(suppl 4):iv26-iv31.
147. Hernández Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. Rheumatology (Oxford). 2013;52(5):944-50.
148. Zink A, Askling J, Dixon WG, Klareskog L, Silman AJ, Symmons DPM. European biologicals registers: methodology, selected results and perspectives. Ann Rheum Dis. 2009;68:1240-6.
149. Lie E, van der Heijde D, Uhlig T, Mikkelsen K, Kalstad S, Kaufmann C, et al. Treatment strategies in patients with rheumatoid arthritis for whom methotrexate monotherapy has failed: data from the NOR-DMARD register. Ann Rheum Dis. 2011;70(12):2103-10.
150. Kvamme MK, Lie E, Kvien TK, Kristiansen IS. Two-year direct and indirect costs for patients with inflammatory rheumatic joint diseases: data from real-life follow-up of patients in the NOR-DMARD registry. Rheumatology (Oxford). 2012;51(9):1618-27.
151. Austad C, Kvien TK, Olsen IC, Uhlig T. Health status has improved more in women than in men with rheumatoid arthritis from 1994 to 2009: results from the Oslo rheumatoid arthritis register. Ann Rheum Dis. 2013 Oct 15; doi: 10.1136/annrheumdis-2013-204014. [Epub ahead of print].
152. Uhlig T, Heiberg T, Mowinckel P, Kvien TK. Rheumatoid arthritis is milder in the new millennium: health status in patients with rheumatoid arthritis 1994–2004. Ann Rheum Dis. 2008;67(12):1710-5.
153. Helsedirektoratet [The Norwegian Directorate of Health]. Insatsstyrt finansiering 2012. [Activity Based Financing]. Oslo: 2011.
154. Den norske legeforening [The Norwegian Medical Association]. Normaltariff for privat spesialistpraksis 2011-2012. [The normal tariff for private specialist practice]. 2012.
155. Helsedirektoratet [The Norwegian Directorate of Health]. Regelverk finansiering av poliklinisk radiologi 2012 - offentlige virksomheter [Regulations for financing outpatient radiology 2012 - public health care]. Oslo 2011.
156. Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. Arthritis Rheum. 2006;55(4):526-31.
157. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006;27(8):861-74.
158. Helsedirektoratet [The Norwegian Directorate of Health]. Helseeffekter i samfunnsøkonomiske analyser [Health effects in economic analyses]. Oslo: 2007.

159. Neovius M, Simard JF, Askling J, Group ftAS. How large are the productivity losses in contemporary patients with RA, and how soon in relation to diagnosis do they develop? Ann Rheum Dis. 2011;70(6):1010-5.
160. Neovius M, Simard JF, Klareskog L, Askling J, Group ftAS. Sick leave and disability pension before and after initiation of antirheumatic therapies in clinical practice. Ann Rheum Dis. 2011;70(8):1407-14.
161. Eriksson JK, Karlsson JA, Bratt J, Petersson IF, van Vollenhoven RF, Ernestam S, et al. Cost-effectiveness of infliximab versus conventional combination treatment in methotrexate-refractory early rheumatoid arthritis: 2-year results of the register-enriched randomised controlled SWEFOT trial. Ann Rheum Dis. 2014 Apr 15; doi:10.1136/annrheumdis-2013-205060 [Epub ahead of print].
162. Karlsson JA, Neovius M, Nilsson J-Å, Petersson IF, Bratt J, van Vollenhoven RF, et al. Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in early rheumatoid arthritis: 2-year quality-of-life results of the randomised, controlled, SWEFOT trial. Ann Rheum Dis. 2013;72(12):1927-33.
163. 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.
164. 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.
165. Rantalaiho V, Kautiainen H, Korpela M, Hannonen P, Kaipiainen-Seppanen O, Mottonen T, et al. Targeted treatment with a combination of traditional DMARDs produces excellent clinical and radiographic long-term outcomes in early rheumatoid arthritis regardless of initial infliximab. The 5-year follow-up results of a randomised clinical trial, the NEO-RACo trial. Ann Rheum Dis. 2013 Aug 1; doi: 10.1136/annrheumdis-2013-203497. [Epub ahead of print].
166. Sokka T, Haugeberg G, Asikainen J, Widding Hansen IJ, Kokko A, Rannio T, et al. Similar clinical outcomes in rheumatoid arthritis with more versus less expensive treatment strategies. Observational data from two rheumatology clinics. Clin Exp Rheumatol. 2013;31(3):409-14.
167. van Vollenhoven RF, Geborek P, Forslind K, Albertsson K, Ernestam S, Petersson IF, et al. Conventional combination treatment versus biological treatment in methotrexate-refractory early rheumatoid arthritis: 2 year follow-up of the randomised, non-blinded, parallel-group Swefot trial. Lancet. 2012;379(9827):1712-20.
168. van Vollenhoven RF, Ernestam S, Geborek P, Petersson IF, Coster L, Waltbrand E, et al. Addition of infliximab compared with addition of sulfasalazine and hydroxychloroquine to methotrexate in patients with early rheumatoid arthritis (Swefot trial): 1-year results of a randomised trial. Lancet. 2009;374(9688):459-66.
169. Boers M, Verhoeven AC, Markusse HM, van de Laar MA, Westhovens R, van Denderen JC, et al. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. Lancet. 1997;350(9074):309-18.
170. de Jong PH, Quax RA, Huisman M, Gerards AH, Feelders RA, de Sonnaville PB, et al. Response to glucocorticoids at 2 weeks predicts the effectiveness of DMARD induction therapy at 3 months: post hoc analyses from the tREACH study. Ann Rheum Dis. 2013;72(10):1659-63.
171. den Uyl D, Ter Wee M, Boers M, Kerstens P, Voskuyl A, Nurmohamed M, et al. A non-inferiority trial of an attenuated combination strategy ('COBRA-light') compared to the original COBRA strategy: clinical results after 26 weeks. Ann Rheum Dis. 2013 April 19; 10.1136/annrheumdis-2012-202818. [Epub ahead of print].
172. 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.
173. de Jong PH, Hazes JM, Han HK, Huisman M, van Zeben D, van der Lubbe PA, et al. Randomised comparison of initial triple DMARD therapy with methotrexate monotherapy in

combination with low-dose glucocorticoid bridging therapy; 1-year data of the tREACH trial. Ann Rheum Dis. 2014 May 1; 10.1136/annrheumdis-2013-204788. [Epub ahead of print].
174. Lillegraven S, Kristiansen IS, Kvien TK. Comparison of utility measures and their relationship with other health status measures in 1041 patients with rheumatoid arthritis. [Report]. Ann Rheum Dis. 2010;69(10):1762-7.
175. Rand-Hendriksen K, Augestad LA, Dahl FA, Kristiansen IS, Stavem K. A shortcut to mean-based time tradeoff tariffs for the EQ-5D? Med Decis Making. 2012;32(4):569-77.
176. Liv Ariane A, Rand-Hendriksen K, Kristiansen IS, Stavem K. Impact of Transformation of Negative Values and Regression Models on Differences Between the UK and US EQ-5D Time Trade-Off Value Sets. Pharmacoeconomics. 2012;30(12):1203-14.
177. Augestad L, Rand-Hendriksen K, Stavem K, Kristiansen I. Time trade-off and attitudes toward euthanasia: implications of using 'death' as an anchor in health state valuation. Qual Life Res. 2013;22(4):705-14.
178. Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. Arthritis Rheum. 2000;43(12):2751-61.
179. Gregersen FA. The impact of ageing on health care expenditures: a study of steepening. The European Journal of Health Economics. 2013 Nov 24; doi: 10.1007/s10198-013-0541-9. [Epub ahead of print].
180. 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. [Miscellaneous Article]. Ann Rheum Dis. 2010;69(4):631-7.
181. Emery P, Kvien TK. Treating rheumatoid arthritis. BMJ. 2007;335(7610):56-7.
182. Putrik P, Ramiro S, Kvien TK, Sokka T, Pavlova M, Uhlig T, et al. Inequities in access to biologic and synthetic DMARDs across 46 European countries. Ann Rheum Dis. 2014;73(1):198-206.
183. Putrik P, Ramiro S, Kvien TK, Sokka T, Uhlig T, Boonen A. Variations in criteria regulating treatment with reimbursed biologic DMARDs across European countries. Are differences related to country's wealth? Ann Rheum Dis. 2013 Aug 12; doi: 10.1136/annrheumdis-2013-203819. [Epub ahead of print].
184. Dorner T, Strand V, Castaneda-Hernandez G, Ferraccioli G, Isaacs JD, Kvien TK, et al. The role of biosimilars in the treatment of rheumatic diseases. Ann Rheum Dis. 2013;72(3):322-8.
185. Schneider CK. Biosimilars in rheumatology: the wind of change. Ann Rheum Dis. 2013;72(3):315-8.
186. Park W, Hrycaj P, Jeka S, Kovalenko V, Lysenko G, Miranda P, et al. A randomised, double-blind, multicentre, parallel-group, prospective study comparing the pharmacokinetics, safety, and efficacy of CT-P13 and innovator infliximab in patients with ankylosing spondylitis: the PLANETAS study. Ann Rheum Dis. 2013;72(10):1605-12.
187. 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;72(10):1613-20.

# 9 PAPER I-III

I

## PAPER I

Kvamme MK, Kristiansen IS, Lie E, Kvien TK

Identification of cutpoints for acceptable health status and important improvement in patient-reported outcomes, in rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis

J Rheumatol. 2010 Jan;37(1):26-31

## PAPER II

Kvamme MK, Lie E, Kvien TK, Kristiansen IS

Two-year direct and indirect costs for patients with inflammatory rheumatic joint diseases: data from real-life follow-up of patients in the NOR-DMARD registry.

Rheumatology (Oxford). 2012 Sep;51(9):1618-27
Oxford University Press

II

## PAPER III

Kvamme MK, Lie E, Uhlig T, Moger TA, Kvien TK, Kristiansen IS
Cost-effectiveness of TNF inhibitors in combination with synthetic DMARDs versus synthetic DMARDs alone in patients with rheumatoid arthritis: a model study based on two longitudinal observational studies

Manuscript, submitted to Rheumatology (Oxford). Status: revise and resubmit.

# Cost-effectiveness of TNF inhibitors versus synthetic disease-modifying antirheumatic drugs in patients with rheumatoid arthritis: a Markov model study based on two longitudinal observational studies

Maria K Kvamme[1,2,3], Elisabeth Lie[1], Till Uhlig[1], Tron A Moger[2], Tore K Kvien[1], Ivar S Kristiansen[2]

Author Affiliations

[1] Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway

[2] Department of Health Management and Health Economics, University of Oslo, Oslo, Norway

[3] Norwegian Knowledge Centre for the Health Services, Oslo, Norway

Correspondence to

Maria Knoph Kvamme, Norwegian Knowledge Centre for the Health Services, PO Box 7004 St. Olavs plass, N-0130 Oslo, Norway; maria.knoph.kvamme@kunnskapssenteret.no; work (+47) 23 25 50 00; mobile (+47) 93 08 32 21 fax (+47) 23 25 50 10



Keywords: “Cost-effectivenes”, “Markov Chain”, “Arthritis, Rheumatoid”, “Antirheumatic Agent”, “Receptors, Tumor Necrosis Factor, Type I”

## Abstract

**Objectives:** The objectives of this study were to estimate the additional costs and health benefits of adding a TNF inhibitor (TNFi) (adalimumab, certolizumab, etanercept, golimumab, infliximab) to a synthetic DMARD (sDMARD), e.g. methotrexate, in patients with rheumatoid arthritis (RA).

**Methods:** We developed the NOrwegian RA (NORA) model as a Markov model simulating 10 years of treatment with either TNFi plus sDMARDs (TNFi-strategy) or sDMARDs alone (synthetic-strategy). Patients in both strategies started in one of seven health states, based on SF-6D. The patients could move to better or worse health states according to transition probabilities. In the TNFi-strategy, patients could stay on TNFi (including switch of TNFi), or switch to non-TNFi-biologics (abatacept, rituximab, tocilizumab), sDMARDs or no DMARD. In the synthetic-strategy, patients remained on sDMARDs. Data from two observational studies were used for the assessment of resource use and utilities in the health states. Health benefits were valued using EQ-5D and SF-6D.

**Results:** The NORA model predicted that ten-year discounted health care costs totalled €124,942 (€475,266 including production losses) for the TNFi-strategy and €65,584 (€436,517) for the synthetic-strategy. The cost per additionally gained quality-adjusted life year (QALY) of adding a TNFi was €92,557 (€60,227 including production losses) using SF-6D and €61,285 (€39,841) using EQ-5D. Including health care costs only, the probability that TNFi-treatment was cost-effective was 90% when using EQ-5D, assuming a Norwegian willingness-to-pay level of €67,300.

**Conclusions:** TNFi-treatment for RA is cost-effective when accounting for production losses. Excluding production losses, TNFi-treatment is cost-effective using EQ-5D, but not SF-6D.

## Background

The effectiveness of biologic disease-modifying antirheumatic drugs (bDMARDs) in the treatment of rheumatoid arthritis (RA) has been documented in a range of randomised controlled clinical trials (RCTs) and systematic reviews (1-5). However, the medical costs of using bDMARDs are at least four to six times higher than using synthetic DMARDs (sDMARDs) (6-8). The importance of evaluating costs in addition to treatment effect is highlighted in the 2010 as well as in the 2013 updated recommendations from the European League Against Rheumatism (EULAR) on the use of DMARDs in RA (9, 10). The 2010 recommendation was underpinned by a systematic literature review of the cost-effectiveness of RA treatments (11).

During the last 10 years, a range of cost-effectiveness analyses/ cost-utility analyses (CEA/CUA) of bDMARDs for patients with RA have been published (12-19). Models have typically been used to simulate health outcomes and costs beyond the relatively short trial periods. The analyses differ in time perspective, cost items, measurement of utility for the calculation of quality-adjusted life years (QALYs) and type of simulation model. The studies have concluded differently with respect to whether bDMARDs are cost-effective.

Previous models have used the Health Assessment Questionnaire Disability Index (HAQ) and/or Disease Activity Score 28 (DAS28) as measures of health outcomes. These outcomes have been converted into utilities that are necessary for cost-utility analyses. Utility is a measure of health related quality of life (HRQoL), measured on a cardinal scale and the utility values are based on preferences. Perfect health has a utility value of 1.0 and death has a utility value of 0.0 (20). These conversion models have used utility instruments such as the EuroQol 5-Dimensions (EQ-5D) and Health Utility Index (HUI) (14-18).

In the majority of previously published cost-effectiveness studies on TNFi, data from clinical trials with limited follow-up time of maximum two years have been used. Exceptions are a study from the UK (19) which used register data with three years of follow-up and a recent Swedish study (15) with almost 10 years of follow-up. Thus to reassess the question is highly relevant when long-term evidence becomes available.

Lack of direct measurement of utility with multi-attribute utility (MAU) instruments such as the EQ-5D as well as limited data on long-term treatment effectiveness, resource use and production losses have been limitations in previous studies (21-23). Thus, no study has assessed the cost-effectiveness using direct measurements of utility within settings of real life long-term data. In the current study, we used up to 10-year follow-up data from the Norwegian DMARD register (NOR-DMARD) (24) for costs and utilities and seven years of follow-up from the Oslo Rheumatoid Arthritis register (ORAR) (25) for patients with conventional standard treatment before bDMARDs were available treatment alternatives.

The objectives of this study were to estimate the incremental costs and health benefits of adding TNFi to sDMARDs in routine care for RA patients compared to treatment with conventional sDMARDs, and to compare the incremental QALYs and cost-effectiveness ratios (ICERs) when using EQ-5D versus SF-6D derived utilities.

# Methods

## Overview of the model

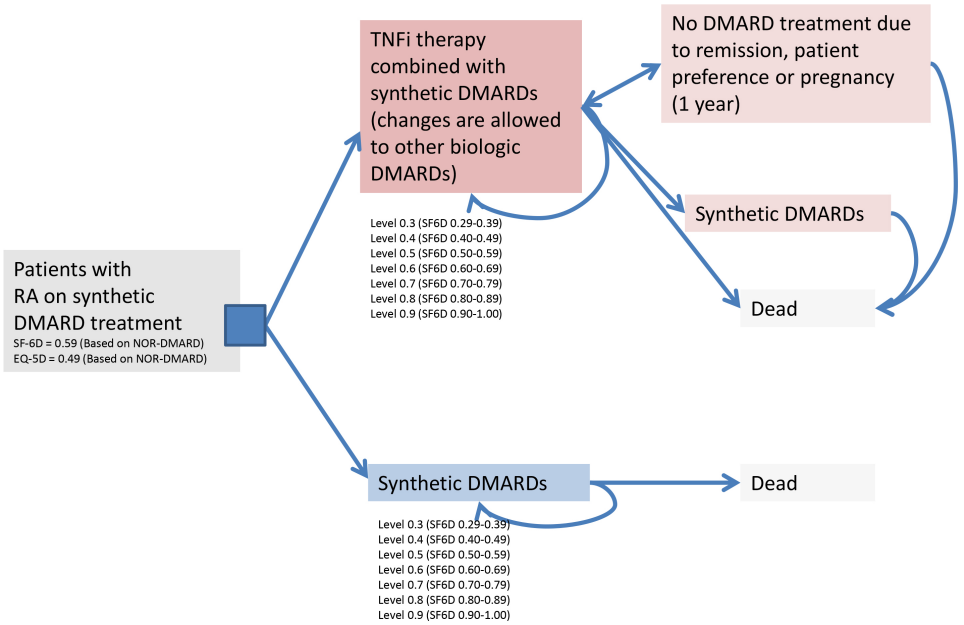
The NOrwegian Rheumatoid Arthritis (NORA) model was developed based on real life data to simulate long-term consequences in terms of health outcomes and costs when using bDMARDs in addition to sDMARDs (Fig. 1). We developed a Markov model that included two main strategies; treatment with sDMARDs in combination with a TNFi (TNFi-strategy) and treatment with sDMARDs alone (synthetic-strategy).

In the model, patients were in different states that were mutually exclusive and collectively exhaustive. There were seven main states in both strategies, based on SF-6D, and in addition death (26). We used a cycle length of 3 months (a quarter of a year), and the time perspective was 10 years. In the TNFi-strategy, patients started with a TNFi (adalimumab, certolizumab, etanercept, golimumab or infliximab) combined with an sDMARD. Patients could remain on the same therapy, switch TNFi, switch to mono TNFi treatment or switch to another bDMARD (abatacept, rituximab or tocilizumab). Patients on all types of bDMARD treatment could be in any of the seven health states and move between them once per quarter. Patients could also move to no DMARD (a specific state in the model), in which they stayed for one year before they moved back to bDMARD treatment. Additionally, patients could move to sDMARDs. Here, they remained in the state for the rest of the analysed period. Finally, patients might die.

In the synthetic-strategy, patients were on sDMARD treatment and did not switch to bDMARD. In line with the TNFi-strategy, patients in the synthetic-strategy could be in any of the seven health states and switch between them and they might die.

**FIG. 1 Overview of the NOrwegian Rheumatoid Arthritis (NORA) model.**

**RA-patients on treatment with conventional sDMARDs at start in the simulation model. They are distributed to two treatment alternatives; TNFi-strategy or continued synthetic-strategy. After a 10-year period, costs and effects of the two strategies are compared.**




**Observational studies: NOR-DMARD and ORAR**

The primary data sources for the NORA model were the observational data from NOR-DMARD (24) and ORAR (25). From 2000, NOR-DMARD recruited patients with inflammatory arthropathies when they started treatment with sDMARDs and/or bDMARDs (24, 27, 28). Clinical assessments were performed at baseline, after 3, 6 and 12 months and then annually until change of DMARD treatment or treatment termination. A new treatment course is similarly evaluated at 3, 6 and 12 months. The TNFi-patients in NOR-DMARD had on average failed two sDMARDs before the initiation of a TNFi (Table 1). In February 2012, NOR-DMARD comprised 7,675 patients of whom 4,079 had a diagnosis of RA. The study was approved by the Norwegian Data Inspectorate and the Regional Ethics Committee of South-Eastern Norway. Patients gave written informed consent before participation.

ORAR was established in 1994 (25). Patients in Oslo (population approximately 450,000 when ORAR was established) with an RA diagnosis were included in the register. We used data from two mail surveys performed in 1994 and 2001 where respondents returned comprehensive questionnaires including SF-6D (based on SF-36). Biological treatment was used by 0% and 3% of all patients in ORAR in 1994 and

2001, respectively. SDMARDs and prednisolone were in the whole ORAR used respectively by 36 % and 41% of patients in 1994 and 48% and 43% in 2001 (29, 30). In total, 412 patients had SF-6D data both in 1994 and in 2001. Of those, 173 were on treatment with an sDMARD in 1994 (Table 1) and six of the 173 were on treatment with a bDMARD in 2001 (Table S1, available at *Rheumatology* Online). We selected the 173 patients who were on sDMARD treatment in 1994 for the synthetic-strategy, thus including patients who continued or discontinued sDMARD treatment and patients who probably would have been offered bDMARDs when these were widely available.

**TABLE 1 Characteristics of RA-patients (mean (range) for continuous, % for categorical variables).**

| **Characteristics** | **NOR-DMARD**<br><br>**Baseline data from patients who started with TNFi plus sDMARD (n=810).** | **ORAR**<br><br>**Baseline data from patients using sDMARD at assessment in 1994 (n=173).** |
|---|---|---|
| Female (%) | 71 | 84 |
| Age at baseline (years) | 52 (18-81) | 54 (25-77) |
| Disease duration (years) | 9 (0-50) | 11 (1-36) |
| Number of previously used sDMARDs | 2.1 (0-7) | NR |
| Number of previously used bDMARDs | 0 | NR |
| SF-6D score | 0.59 (0.31-0.95) | 0.63 (0.39-1.0) |
| EQ-5D score | 0.49 ( -0.248-1.0) (n=456) | NR |
| MHAQ score | 0.71 (0-2.75) | 0.52 (0-2) |
| DAS28 level | 2.32 (0-3) | NR |
| *Level of participation in the workforce* | | NR |
| Retired due to age (%) | 12 | NR |
| On disability pension or rehabilitation benefits (%) | 33 | NR |
| On sick leave (%) | 14 | NR |
| Fully employed (%) | 21 | NR |
| Working part time (%) | 14 | NR |
| Other (students, unemployed etc.) (%) | 6 | NR |

NOR-DMARD, the Norwegian disease-modifying antirheumatic drug register; ORAR, the Oslo rheumatoid arthritis register; sDMARDs, synthetic disease-modifying antirheumatic drugs; bDMARDs, biologic disease-modifying antirheumatic drugs; SF-6D, Short Form-6 Dimensions (0.296-1.0, 1.0 best); EQ-5D, EuroQol-5 Dimensions (-0.594 to 1.0, 1.0 best); MHAQ, Modified Health Assessment Questionnaire (0-3, 0 best); DAS28, Disease Activity Score 28 (0 to >5.1, 0 best); NR, not reported.

## Markov states and utilities in the states

In Markov models, patients move between predefined health states at the end of each time cycle the model runs. We defined the health states on the basis of the SF-6D which is derived from the Medical Outcomes Study 36-item Short-Form Health Survey (SF-36) (31). We chose to use SF-6D for the division into health states because it can be used to generate QALYs and because SF-36 data were available at each assessment both in NOR-DMARD and ORAR.

We translated SF-36 scores to utilities by means of a validated U.K. algorithm (26). The SF-6D algorithm implies that no health state is below 0.296 except for death, (32) and consequently we divided the patients into seven groups; ≥0.296-0.4, >0.4-0.5, etc. up to 1.0 (Fig. 1 and Table 2).

At start of the simulation, the two simulation cohorts had the same characteristics with respect to age, gender and HRQoL distributions. Patients in both strategies were distributed across Markov-states according to findings in NOR-DMARD (n=810) (Table 1-2).

**TABLE 2 Distribution of patients in health states (SF-6D) at initiation of treatment with TNFi plus sDMARDs (NOR-DMARD)**

| | **NOR-DMARD, n=810** | |
|---|---|---|
| **Health states** (Range 0.296-1.0) | Frequency | % |
| ≥0.296 and ≤ 0.4 | 51 | 6 |
| >0.4 and ≤ 0.5 | 109 | 13 |
| >0.5 and ≤ 0.6 | 312 | 39 |
| >0.6 and ≤ 0.7 | 193 | 24 |
| >0.7 and ≤ 0.8 | 84 | 10 |
| >0.7 and ≤ 0.9 | 57 | 7 |
| >0.9 and ≤ 1.0 | 4 | 0 |
| Total | 810 | |

NOR-DMARD, the Norwegian disease-modifying antirheumatic drug register; sDMARDs, synthetic disease-modifying antirheumatic drugs; SF-6D, Short Form-6 Dimensions (0.296-1.0, 1.0 best); TNFi, tumour necrosis factor inhibitor.

We used SF-6D- as well as EQ-5D-utilities in the analyses (26, 33). The EQ-5D was included in NOR-DMARD in 2006 which means that we had fewer patients with EQ-5D data (n=456) than with SF-6D (n=810). The mean SF-6D/EQ-5D score for each state at treatment start from NOR-DMARD (n=810)/ (n=456) were used for the utility estimates in the corresponding states in the model (Table 3). The change in utility over time in the TNFi-strategy was based on NOR-DMARD data for the 810 patients who started treatment with a TNFi plus methotrexate or other sDMARDs during 2001-2011. Patients who switched to sDMARD treatment (n=225) had a mean SF-6D utility value of 0.635 (95% CI:0.621,0.650) (EQ-5D utility value of 0.553 (95% CI:0.505,0.601),n=105). Patients who discontinued the follow-up due to remission, patient preference, or planned pregnancy (n=64) had at exclusion a mean SF-6D utility of 0.718 (95% CI:0.684,0.752) (EQ-5D 0.745 (95% CI:0.622,0.869), n=29).

In the synthetic-strategy, patients initially had the same utility distribution as in the bDMARD strategy, but the changes in utility level over time were based on the 173 patients in ORAR who were on sDMARD treatment in 1994 and had SF-6D data both in 1994 and in 2001. Their mean SF-6D score was 0.631 (95% CI:0.613,0.649) in 1994 and 0.634 (95% CI:0.614,0.655) in 2001, implying no mean change in utility over these seven years. Excluding the six ORAR patients using bDMARDs in 2001 did not alter the change in utility. There was no EQ-5D data in ORAR. However, since the SF-6D data did not indicate any deterioration over the seven-year period, we assumed no deterioration in utility scores for patients in the synthetic strategy, also when using EQ-5D as outcome measure.

**TABLE 3 Baseline SF-6D and EQ-5D mean utility values across health states in NOR-DMARD.**

| Health states (Range 0.296-1.0) | Mean SF-6D (n=810) | Mean EQ-5D (n=456) |
|---|---|---|
| ≥0.296 and ≤ 0.4 | 0.37 | -0.03 |
| >0.4 and ≤ 0.5 | 0.46 | 0.23 |
| >0.5 and ≤ 0.6 | 0.56 | 0.45 |
| >0.6 and ≤ 0.7 | 0.64 | 0.56 |
| >0.7 and ≤ 0.8 | 0.74 | 0.75 |
| >0.7 and ≤ 0.9 | 0.85 | 0.79 |
| >0.9 and ≤ 1.0 | 0.94 | 0.84 |
| Total (95% CI) | 0.593 (0.585, 0.601) | 0.490 (0.462, 0.518) |

NOR-DMARD the Norwegian disease-modifying antirheumatic drug register; SF-6D Short Form-6 Dimensions (0.296-1.0, 1.0 best); EQ-5D EuroQol-5 Dimensions (-0.594 to 1.0, 1.0 best).

## Transitions between health states

Patients may over time move to other health states according to transition probabilities generated from NOR-DMARD and ORAR data (24). The probability for a patient to change health state was calculated by ratios, taken from the number of transitions from a given state and to another (Table S2A-D, available at *Rheumatology* Online).

The probabilities of dying were based on mortality tables from Statistics Norway. The mortality rates were adjusted for RA, using a published standardized mortality ratio (SMR) of 1.54 (34). We assumed that mortality risk was equal in the two strategies.

Patients on TNFi may for different reasons discontinue treatment. The reasons for treatment termination, as registered in NOR-DMARD are presented in Table S3, available at *Rheumatology* Online. Since patients had been included continuously in NOR-DMARD since 2000, only a minority of patients had reached 10-years of follow-up.

Patients may change treatment in the TNFi-strategy, according to NOR-DMARD data. An overview of number of patients with registry data who were on treatment with a TNFi plus sDMARDs, TNFi monotherapy, non-TNFi biologic or sDMARD is presented in Table S4. Detailed treatment types for patients in the TNFi-strategy are presented in Table S5, available at *Rheumatology* Online. Details about

drug treatment for patients in the synthetic-strategy are presented in Table S1, available at *Rheumatology* Online. Patients included in the synthetic-strategy could in real life have changed sDMARD treatment or discontinued treatment. All are included in the synthetic-strategy; labelled Synthetic DMARDs (Fig. 1).

## Costs

Cost estimates were based on identification, quantification and valuation of resource use and level of labour force participation. All cost estimates were based on 2012 unit prices. Prices were exclusive of value-added tax. Patient co-payments and travel expenses were included.

Unit costs (except drug costs) are presented in Table S6, available at *Rheumatology* Online. Details of the cost calculations from NOR-DMARD are presented in Text S1, available at *Rheumatology* Online, and in a previous publication (6).

The costs in each state were the mean costs of patients in the corresponding health state from the NOR-DMARD register. Patients in the TNFi-strategy received costs in each cycle that corresponded to their health state and treatment type. Patients in the synthetic-strategy were assigned to the costs of patients in NOR-DMARD on sDMARD treatment, recorded before the same patients started treatment with a TNFi. Details of the cost components in the different health states are presented Table S7, available at *Rheumatology* Online.

## Discounting

Costs and effects were discounted at an annual rate of 4 % as recommended in Norway (35).

## Uncertainty

We performed deterministic sensitivity analyses of SF-6D versus EQ-5D, different willingness-to-pay (WTP) thresholds (online supplementary table S6) and no versus four percent discount rate (not shown). Willingness-to-pay is "a method of measuring the value an individual places on a good, service, or reduction in the risk of death and illness by estimating the maximum dollar amount an individual would pay in order to obtain the good, service or risk reduction" (36 p. 335). In our analyses, the WTP-threshold reflects the maximum willingness-to-pay for a QALY from a societal perspective. We illustrated the results with WTP-thresholds from approximately €50,000 to €100,000. Probabilistic sensitivity analyses were performed for probabilities of changing health states, health care costs, total costs including production losses, and utility values in health states. Uncertainty in the cost estimates was expressed by gamma distributions, utility decrements by gamma distributions, and transition probabilities by Dirichlet distributions. The latter distribution was used because it is the multivariate version of the beta distribution which is recommended in a probability parameter where the data are binomial (37). Standard errors and means from the empirical cost-and utility data were used for parameters in the gamma distributions and numbers of actual transitions were used for parameters in the Dirichlet distributions. All distributions are presented in Table S8, available at *Rheumatology* Online. The parameter uncertainty was assessed in Monte Carlo simulations. 1000 simulations were run for each

combination of costs and effects. Probabilities for the estimated results to be below a chosen WTP-threshold are presented in Fig. S1 and Table S9, available at *Rheumatology* Online.

Data were analysed using Stata MP11, College Station, TX, USA and the model was developed in TreeAge Pro 2012 and 2013.

## Results

The estimated discounted health care costs were €124,942 in the TNFi-strategy over the 10-year period and €65,584 in the synthetic-strategy. The costs were €475,266 and €436,517 respectively, when production losses were included (Table 4). These results imply that the incremental cost per QALY was €92,557 using SF-6D and €61,285 using EQ-5D, excluding production losses. Including production losses, the ICERs diminished to €60,227 using SF-6D and €39,841 using EQ-5D (Table 4).

**TABLE 4 Cost-effectiveness results for a 10-year period**

| Type of cost | Type of utility instru-ment | Strategy | Costs in €* | QALYs* | Incre-mental costs in €* | Incre-mental QALYs* | ICER €* |
|---|---|---|---|---|---|---|---|
| Health care | SF-6D | Synthetic | 65,584 | 4.78 | | | |
| | | TNFi | 124,942 | 5.42 | 59,358 | 0.64 | 92,557 |
| | EQ-5D | Synthetic | 65,593 | 3.82 | | | |
| | | TNFi | 124,941 | 4.79 | 59,348 | 0.97 | 61,285 |
| Health care + production losses | SF-6D | Synthetic | 436,517 | 4.78 | | | |
| | | TNFi | 475,266 | 5.42 | 38,749 | 0.64 | 60,227 |
| | EQ-5D | Synthetic | 438,012 | 3.82 | | | |
| | | TNFi | 476,653 | 4.79 | 38,641 | 0.97 | 39,841 |

QALYs, quality-adjusted life-years; ICER, Incremental cost-effectiveness ratio
*Costs and effects are discounted at 4 %

Including health care costs and using SF-6D, the probability was zero that TNFi treatment was cost-effective at a WTP-threshold of €67,300. This value that has been proposed by the Norwegian Directorate of Health as a "best estimate" for the value of a statistical life with full health (38). If EQ-5D was used, the probability that TNFi-treatment was cost-effective was 90% at the same WTP-threshold. When also production losses were included, TNFi-treatment was cost-effective, irrespective of type of utility instrument used. Using SF-6D, the probability that TNFi-treatment was cost-effective was 89%. EQ-5D utilities yielded a probability of 100% for the same strategy, already at a WTP-threshold of NOK400,000 (€53,800) (Table S9, available at *Rheumatology* Online).

Incremental cost-effectiveness scatter plots illustrating the incremental costs with the TNFi-strategy plotted on the y-axis and the incremental QALYs on the x-axis and cost-effectiveness acceptability curves for the base case analysis are presented in Fig. 2A-D. Cost-effectiveness acceptability curves for the base case analysis (health care costs and EQ-5D utilities) is presented in Fig. S1, available at *Rheumatology* Online.

**FIG. 2A-D**

**Fig. 2A Incremental cost-effectiveness scatter plot. Health care costs in € and EQ-5D utilities.**

**The incremental costs with the TNFi-strategy are plotted on the y-axis and the incremental QALYs with the TNFi-strategy are plotted on the x-axis. A WTP-threshold of approximately €67,000 is represented by the green line. In 90% of the simulations, the incremental cost-effectiveness ratio was below this threshold (i.e. to the south-east of the grey line).**

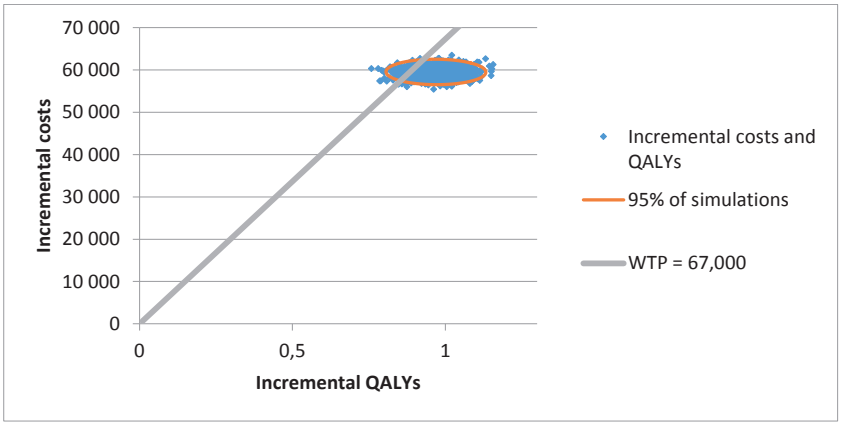



**Fig. 2B Incremental cost-effectiveness scatter plot. Health care costs in € and SF-6D utilities.**

**The incremental costs and QALYs with the TNFi-strategy are plotted in the scatter plot. A WTP-threshold of approximately €67,000 is represented by the grey line.**

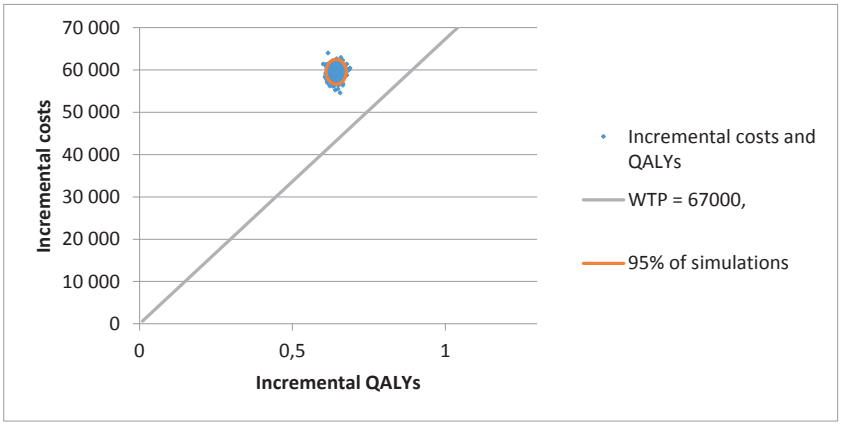


**Fig. 2C Incremental cost-effectiveness scatter plot. Health care costs and production losses in € and EQ-5D utilities.**

**The incremental costs and QALYs with the TNFi-strategy are plotted in the scatter plot. A WTP-threshold of approximately €67,000 is represented by the grey line.**

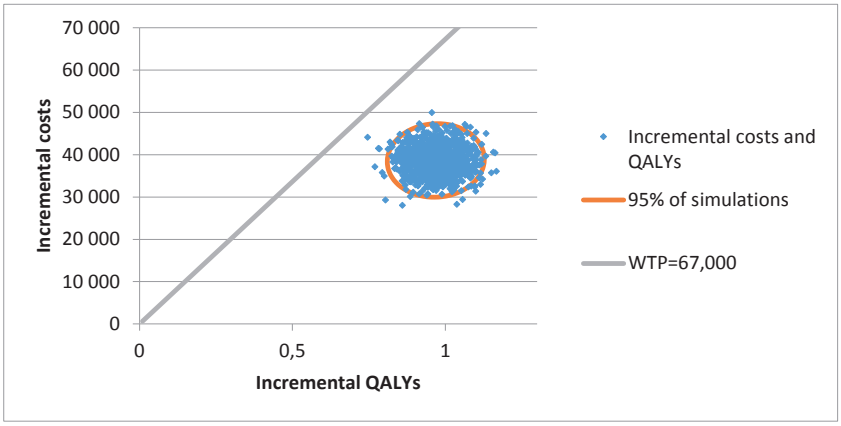
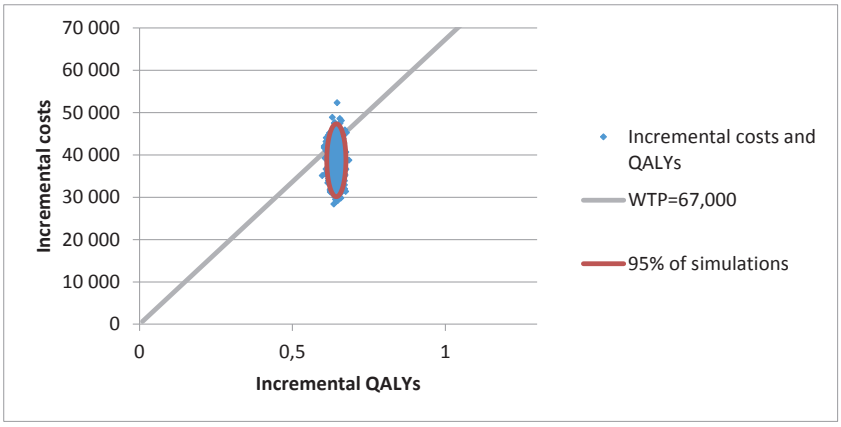



**Fig. 2D Incremental cost-effectiveness scatter plot. Health care costs and production losses in € and SF-6D utilities.**

**The incremental costs and QALYs with the TNFi-strategy are plotted in the scatter plot. A WTP-threshold of approximately €67,000 is represented by the grey line.**



In our setting of routine care for patients with RA, TNFi in combination with sDMARDs were cost-effective compared to sDMARD alone when including production losses.T This conclusion is based on health benefits valued directly with both SF-6D and EQ-5D. However, SF-6D values the health benefits lower than EQ-5D. Disregarding production losses, sDMARDs were cost-effective only when using EQ-5D.

A recent review indicate that studies examining the cost-effectiveness of bDMARDs versus sDMARDs are inconclusive in models based on clinical trials (21). The review reported eight studies which evaluated the cost-effectiveness of adding a biologic therapy (including adalimumab, etanercept and infliximab) into a DMARD sequence versus a sDMARD sequence in patients who had failed two or more sDMARDs (including leflunomide, MTX, sulfasalazine, combination therapies and sDMARD sequences). The biologic alternative was cost-effective in 14 of 35 comparisons, applying a WTP-threshold of Can $100,000 (€74,465) per QALY. The ICER values ranged from €33,500 to €456,200. Only one of the studies had a societal perspective, the others had a payer perspective (21, 39). In our study, all combinations except for health care costs and SF-6D utilities generated ICERs below a WTP-threshold of €74,465 per QALY. NOR-DMARD included prescription- and infusion data on all the TNFis (adalimumab, certolizumab, etanercept, golimumab, infliximab) and we assumed equal effectiveness of all the TNFis on a group level which is in line with current EULAR recommendations (9, 10).

Compared to a Swedish cost-effectiveness study which also used registry data, our ICERs were higher (15, 40). The difference might partly be explained by the assumptions made regarding the change in utility over time for RA-patients in the comparator arm. We assumed that patients on sDMARD treatment had no deterioration in utility level over a period of 10-years, while the Swedish study assumed that the alternative to bDMARDs entailed natural progression and an annual increase in HAQ-score of 0.065 (15). This level of progression is higher than that reported in other studies in the pre-biologic era. For example, Wolfe et al and Scott et al found an annual HAQ-score progression of 0.03 (41, 42). Importantly, the data from our sDMARD arm are from 1994 to 2001, i.e. mainly before bDMARDs were introduced in 1999.

To our knowledge, our RA model is the first in which the health states were defined on the basis of HRQoL data directly measured as utilities. Because both of our observational studies, NOR-DMARD and ORAR, included the SF-36 for calculation of SF-6D utilities, we could base the Markov model directly on these utility data. Previous models have used HAQ-score or other proxies; requiring translation of the proxy into utilities, which will inevitably increase uncertainty. NOR-DMARD included EQ-5D in 2006, which gave us the opportunity to compare utilities from two separate generic utility instruments. We had up to 10 years of follow-up for patients in the TNFi-strategy and seven years in ORAR for patients in the synthetic-strategy. An additional strength was the detailed cost data in NOR-DMARD (6), which allowed estimation of actual costs of pharmaceuticals as well as hospital care, physiotherapy, rehabilitation and work absenteeism. Register data represent real-life treatment and resource use as compared to data from RCTs, which tend to have protocol driven costs and selected patient groups.

This model study has some weaknesses that may threaten the validity of the results. First, NOR-DMARD typically recruited patients when they had a deterioration of the disease while patients in the ORAR were measured at random points in the disease development. This may imply a regression to the mean (43)

bias in favour of bDMARDs. Second, the ORAR data were mainly collected before the introduction of bDMARDs. The treatment of RA has changed and is currently more intensive. This more aggressive treatment approach might underestimate the effectiveness of the synthetic-strategy. Third, the ORAR data imply that patients on sDMARDs have a constant HRQoL over 7 years, while models conventionally assume that HRQoL deteriorates over time. The stable level may be explained by more active use of sDMARDs, faster adjustment in therapies as well as better patient care in general, even before the introduction of bDMARDs. Fourth, the sDMARD costs were based on the last 3 months of treatment prior to bDMARD treatment in the NOR-DMARD register. In this period, patients' presumably had higher costs than they generally have. The first and second weakness will bias the ICER down while the third will bias the ICER up and the fourth will bias the ICER down. The last impact may be counteracted, however, because treatment hospital costs increase with age (44).

In conclusion, this study indicates that the estimated cost-effectiveness of bDMARDs in RA depends on the choice of utility instrument and whether or not production losses are included in the analyses. We suggest that the results from the NORA model should be validated in other studies, including long-term randomized controlled clinical trials that have included SF-6D and/or EQ-5D as outcome measures. Further, the discrepancy in the cost-effectiveness findings based on choice of utility instrument should encourage payers of expensive medications to reach consensus on use of utility instruments in cost–effectiveness analyses.

Key messages

Long-term real life observational data indicate that TNFi-treatment for RA may be cost-effective.

Choice of utility instrument influences the results of cost-effectiveness analyses.

Acknowledgements

We are grateful toward Elisabeth Orskaug and Ingunn Friede Tvete at the Norwegian Computing Centre for inputs on model structure and estimation of transition probabilities. Cathrine Austad at the Department of Rheumatology, Diakonhjemmet Hospital, provided valuable help in preparation and extraction of ORAR data.

Conflict of interests

MKK has received travel/accommodations/course expenses from Pfizer. EL has received consulting and/or speaker honoraria from Abbvie, Bristol-Myers Squibb, Hospira, Pfizer, Roche and UCB. TU has received consulting and/or speaker honoraria from Abbvie, Bristol-Myers Squibb, MSD, Pfizer, Roche and UCB. TAM has nothing to declare. TKK has received fees for speaking and/or consulting from AbbVie, BMS, Celltrion, Eli Lilly, Hospira, MSD/Schering-Plough, Orion Pharma, Pfizer/Wyeth, Roche, UCB and received research funding to the Diakonhjemmet Hospital from AbbVie, BMS, MSD/Schering-Plough,

Pfizer/Wyeth, Roche and UCB. ISK has received gifts, travel funds, honorariums, consultancy fees or salary from a wide range of public institutions, not for profit organisations or for profit organisations that may have an interest in DMARDs.

Funding

This work was supported by the South-Eastern Norway Regional Health Authority. The NOR-DMARD study has previously received grant support from AbbVie, Amgen, Wyeth/Pfizer, Aventis, MSD, Schering-Plough/Centocor, Bristol-Myers Squibb, UCB, Roche and the Norwegian Directorate for Health and Social Affairs.

## References


1. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. The Lancet. 2004;363:675-81.

2. Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M, et al. Infliximab (chimeric anti-tumour necrosis factor@monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. The Lancet. 1999;354:1932-9.

3. Breedveld F, Weisman M, Kavanaugh A, Cohen S, 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 & rheumatism. 2006;54:26-37.

4. Knevel R, Schoels M, Huizinga TWJ, Aletaha D, Burmester GR, Combe B, et al. Current evidence for a strategic approach to the management of rheumatoid arthritis with disease-modifying antirheumatic drugs: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis. Annals of the Rheumatic Diseases. 2010;69:987-94.

5. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. Biologics for rheumatoid arthritis: an overview of Cochrane reviews. The Cochrane database of systematic reviews. 2009:Cd007848.

6. Kvamme MK, Lie E, Kvien TK, Kristiansen IS. Two-year direct and indirect costs for patients with inflammatory rheumatic joint diseases: data from real-life follow-up of patients in the NOR-DMARD registry. Rheumatology. 2012;51:1618-27.

7. Fautrel B, Verstappen SM, Boonen A. Economic consequences and potential benefits. BestPractResClinRheumatol. 2011;25:607-24.

8. Van den Hout WB, Goekoop-Ruiterman YPM, Allaart CF, De Vries-Bouwstra JK, M. Hazes JM, Kerstens PJSM, et al. Cost-utility analysis of treatment strategies in patients with recent-onset rheumatoid arthritis. Arthritis Care Res. 2009;61:291-9.

9. Smolen JS, Landewé 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. Annals of the Rheumatic Diseases. 2010;69:964-75.

10. Smolen JS, Landewé 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. Annals of the Rheumatic Diseases. 2014;73:492-509.

11. Schoels M, Wong J, Scott DL, Zink A, Richards P, Landewé R, et al. Economic aspects of treatment options in rheumatoid arthritis: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis. Annals of the Rheumatic Diseases. 2010;69:995-1003.

12. Wong JB, Singh G, Kavanaugh A. Estimating the cost-effectiveness of 54 weeks of infliximab for rheumatoid arthritis. The American Journal of Medicine. 2002;113:400-8.

13. Kobelt G, Jönsson L, Young A, Eberhardt K. The cost–effectiveness of infliximab (Remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study. Rheumatology. 2003;42:326-35.

14. Kobelt G, Lindgren P, Singh A, Klareskog L. Cost effectiveness of etanercept (Enbrel) in combination with methotrexate in the treatment of active rheumatoid arthritis based on the TEMPO trial. Annals of the Rheumatic Diseases. 2005;64:1174-9.

15. Lekander I, Borgström F, Svarvar P, Ljung T, Carli C, van Vollenhoven RF. Cost-effectiveness of real-world infliximab use in patients with rheumatoid arthritis in Sweden. International Journal of Technology Assessment in Health Care. 2010;26:54-61.

16. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumar necrosis factor in rheumatoid arthritis. Health Technology Assessment. 2004;8.

17. Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. Rheumatology. 2004;43:62-72.

18. Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. Annals of the Rheumatic Diseases. 2005;64:995-1002.

19. Brennan A, Bansback N, Nixon R, Madan J, Harrison M, Watson K, et al. Modelling the cost effectiveness of TNF-α antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. Rheumatology. 2007;46:1345-54.

20. Torrance GW. Utility approach to measuring health-related quality of life. J Chronic Dis. 1987;40:593-600.

21. van der Velde G, Pham B, Machado M, Ieraci L, Witteman W, Bombardier C, et al. Cost-effectiveness of biologic response modifiers compared to disease-modifying antirheumatic drugs for rheumatoid arthritis: A systematic review. Arthritis Care Res. 2011;63:65-78.

22. Tosh J, Brennan A, Wailoo A, Bansback N. The Sheffield rheumatoid arthritis health economic model. Rheumatology. 2011;50:iv26-iv31.

23. Hernández Alava M, Wailoo A, Wolfe F, Michaud K. The relationship between EQ-5D, HAQ and pain in patients with rheumatoid arthritis. Rheumatology. 2013;52:944-50.

24. Kvien TK, Heiberg MS, Lie E, Kaufmann C, Mikkelsen K, Nordvåg B-Y, et al. A Norwegian DMARD register: Prescriptions og DMARDs and biological agents to patients with inflammatory rheumatic diseases. Clin Exp Rheumatol. 2005;23:s188-94.

25. Kvien TK, Glennas A, Knudsrod OG, Smedstad LM, Mowinckel P, Forre O. The prevalence and severity of rheumatoid arthritis in Oslo. Results from a county register and population survey. Scand J Rheumatol. 1997;26:412-8.

26. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics. 2002;21:271-92.

27. Zink A, Askling J, Dixon WG, Klareskog L, Silman AJ, Symmons DPM. European biologicals registers: methodology, selected results and perspectives. Annals of the Rheumatic Diseases. 2009;68:1240-6.

28. Lie E, van der Heijde D, Uhlig T, Mikkelsen K, Kalstad S, Kaufmann C, et al. Treatment strategies in patients with rheumatoid arthritis for whom methotrexate monotherapy has failed: data from the NOR-DMARD register. Annals of the Rheumatic Diseases. 2011;70:2103-10.

29. Uhlig T, Heiberg T, Mowinckel P, Kvien TK. Rheumatoid arthritis is milder in the new millennium: health status in patients with rheumatoid arthritis 1994–2004. Annals of the Rheumatic Diseases. 2008;67:1710-5.

30. Austad C, Kvien TK, Olsen IC, Uhlig T. Health status has improved more in women than in men with rheumatoid arthritis from 1994 to 2009: results from the Oslo rheumatoid arthritis register. Annals of the Rheumatic Diseases. 2013 published on 15 October 2013. doi:10.1136/annrheumdis-2013-204014.

31. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. Medical Care. 1992;30:473-83.

32. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. Health Economics. 2004;13:873-85.

33. Dolan P. Modeling Valuations for EuroQol Health States. Medical Care. 1997;35:1095-108.

34. Sokka T, Abelson B, Pincus T. Mortality in rheumatoid arthritis: 2008 update. Clinical and Experimental Rheumatology. 2008;26:0035-61.

35. Economic evaluation of healthcare interventions - a guide. The Norwegian Directorate of Health; 2012.

36. Brazier J, Ratcliffe J, Salomon J, A, Tsuchiya A. Measuring and Valuing Health Benefits for Economic Evaluation. New York: Oxford University Press, Inc.; 2007.

37. Briggs A, Claxton K, Schulper M. Decision Modelling for Health Economic Evaluation: Oxford University Press; 2006.

38. Sælensminde K. Health effects in economic analyses. Oslo: The Norwegian Directorate of Health, 2007 IS-1435.

39. Welsing PMJ, Severens JL, Hartman M, van Riel PLCM, Laan RFJM. Modeling the 5-year cost effectiveness of treatment strategies including tumor necrosis factor-blocking agents and leflunomide for treating rheumatoid arthritis in the Netherlands. Arthritis and Rheumatism. 2004;51.

40. Kobelt G, Lindgren P, Lindroth Y, Jacobson L, Eberhardt K. Modelling the effect of function and disease activity on costs and quality of life in rheumatoid arthritis. Rheumatology. 2005;44:1169-75.

41. Wolfe F. A reappraisal of HAQ disability in rheumatoid arthritis. Arthritis Rheum. 2000;43:2751-61.

42. Scott DL, Pugner K, Kaarela K, Doyle DV, Woolf A, Holmes J, et al. The links between joint damage and disability in rheumatoid arthritis. Rheumatology. 2000;39:122-32.

43. Upton G, Cook I. A Dictionary of Statistics. 2 rev. ed. Published online: Oxford University Press; 2008.

44. Gregersen F. The impact of ageing on health care expenditures: a study of steepening. The European Journal of Health Economics. 2013 published on 24 November 2013. doi:10.1007/s10198-013-0541-9.

## Supplements Paper III

**Supplementary table 1. Treatment types in RA patients from ORAR.**

| **DMARDs** | Used in | | | |
|---|---|---|---|---|
| | 1994 | | 2001 | |
| **sDMARDs** | n | % | n | %[1] |
| Azathioprine | 3 | 2 % | 4 | 2 % |
| Cyclosporine | 6 | 3 % | 2 | 1 % |
| Hydroxychloroquine | 10 | 6 % | 3 | 2 % |
| Hydroxychloroquine + intramuscular gold | 1 | 1 % | | |
| Hydroxychloroquine + sulphasalazine | | | 2 | 1 % |
| Intramuscular gold | 23 | 13 % | 9 | 5 % |
| Leflunomide | | | 6 | 3 % |
| Methorexate | 83 | 48 % | 65 | 38 % |
| Methorexate + cyclosporine | | | 3 | 2 % |
| Methorexate + cyclosporine + hydroxychloroquine | | | 1 | 1 % |
| Methotrexate + hydroxychloroquine | 2 | 1 % | 4 | 2 % |
| Methotrexate + hydroxychloroquine + sulphasalazine | | | 5 | 3 % |
| Methotrexate + oral gold | 1 | 1 % | | |
| Methotrexate + sulphasalazine | 1 | 1 % | 3 | 2 % |
| Oral gold | 7 | 4 % | 1 | 1 % |
| Other | 6 | 3 % | | |
| Sulfasalazine | 30 | 17 % | 8 | 5 % |
| **bDMARDs** | | | 6 | 3 % |
| Total | 173 | 100 % | 122 | 67 % |

[1] Percentage of the 173 who used sDMARDs in 1994

**Supplementary tables 2A-D. The transition probabilities presented below are based on data in NOR-DMARD, from patients starting treatment with a TNFi + sDMARD.**

The probabilities of a change of health state during the first two quarters, *i.e.* the probability of a transition from a health state to all other health states from the first to the second quarter (**table S2A**) and from the second to the third quarter (**table S2B**) were calculated separately since we had quarterly updates from the register for these transitions. Thereafter, the yearly probabilities of a change, *i.e* from quarter 4, 8, 12, 16, 20, 24, 28, 32, 36 and 40 were calculated separately since the probability of a change was higher when the yearly assessments were scheduled (**table S2C**). The probabilities of changing between the other quarters were estimated separately since these probabilities were lower than the yearly changes (**table S2D)**. The probabilities of a switch to sDMARDs and to no DMARD (due to remission, pregnancy or the patient's desire) were also calculated by ratios directly from the data in NOR-DMARD.

**Table S2A**. Patients in the TNFi-strategy according to Markov states in 1st and 2nd quarter, n=632[2].

| Markov-state in 1st quarter | Markov-state in 2nd quarter | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Synthetic** | **No DMARD** | **0.3-0.4** | **0.4-0.5** | **0.5-0.6** | **0.6-0.7** | **0.7-0.8** | **0.8-0.9** | **0.9-1.0** | ***Total*** |
| **0.3-0.4** | 2 | 1 | 2 | 8 | 18 | 6 | 2 | 2 | 1 | 41 |
| **0.4-0.5** | 0 | 0 | 5 | 13 | 33 | 25 | 5 | 3 | 0 | 84 |
| **0.5-0.6** | 8 | 0 | 1 | 9 | 79 | 85 | 34 | 21 | 5 | 242 |
| **0.6-0.7** | 5 | 0 | 0 | 0 | 20 | 57 | 33 | 33 | 7 | 155 |
| **0.7-0.8** | 3 | 0 | 0 | 1 | 2 | 16 | 20 | 21 | 3 | 66 |
| **0.8-0.9** | 0 | 0 | 0 | 1 | 1 | 4 | 4 | 25 | 7 | 42 |
| **0.9-1.0** | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| ***Total*** | 18 | 2 | 8 | 32 | 153 | 193 | 98 | 105 | 23 | 632 |

**Table S2B**. Number of patients with transitions from Markov-states to Markov-states, from 2nd to 3rd quarter of treatment with a TNFi in combination with sDMARD.

| Markov-state in 2nd quarter | Markov-state in 3rd quarter | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Synthetic** | **No DMARD** | **0.3-0.4** | **0.4-0.5** | **0.5-0.6** | **0.6-0.7** | **0.7-0.8** | **0.8-0.9** | **0.9-1.0** | ***Total*** |
| **0.3-0.4** | 0 | 0 | 3 | 1 | 1 | 0 | 0 | 0 | 0 | 5 |
| **0.4-0.5** | 1 | 0 | 3 | 11 | 11 | 2 | 0 | 0 | 0 | 28 |
| **0.5-0.6** | 8 | 0 | 3 | 17 | 59 | 31 | 8 | 6 | 0 | 132 |
| **0.6-0.7** | 2 | 0 | 1 | 2 | 31 | 77 | 27 | 19 | 2 | 161 |
| **0.7-0.8** | 1 | 0 | 0 | 1 | 9 | 23 | 24 | 21 | 2 | 81 |
| **0.8-0.9** | 2 | 0 | 0 | 0 | 4 | 13 | 16 | 37 | 11 | 83 |
| **0.9-1.0** | 1 | 1 | 0 | 0 | 0 | 2 | 1 | 5 | 6 | 16 |
| ***Total*** | 15 | 1 | 10 | 32 | 115 | 148 | 76 | 88 | 21 | 506 |

[2] At the first visit there were 810 patients with available SF-6D data, and at the second visit the number was 632. The reasons for the drop in number of patients were that patients had left the study for different reasons, missing SF-6D data, and that patients had only just started the treatment and not yet reached the second visit.

**Table S2C**. Number of patients with transitions from Markov-states to Markov-states, based on yearly assessments for patients on treatment with a TNFi in combination with sDMARD. These transitions are taking place in one out of four quarters in the simulation.

| Markov-state previous | Markov-state in subsequent quarter | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Synthetic** | **No DMARD** | **0.3-0.4** | **0.4-0.5** | **0.5-0.6** | **0.6-0.7** | **0.7-0.8** | **0.8-0.9** | **0.9-1.0** | ***Total*** |
| **0.3-0.4** | 1 | 0 | 10 | 5 | 5 | 2 | 1 | 0 | 0 | 24 |
| **0.4-0.5** | 3 | 0 | 7 | 36 | 23 | 11 | 2 | 1 | 0 | 83 |
| **0.5-0.6** | 7 | 3 | 7 | 15 | 256 | 79 | 21 | 13 | 0 | 401 |
| **0.6-0.7** | 10 | 2 | 3 | 10 | 93 | 278 | 69 | 44 | 3 | 512 |
| **0.7-0.8** | 4 | 4 | 0 | 3 | 23 | 68 | 138 | 61 | 5 | 306 |
| **0.8-0.9** | 3 | 3 | 1 | 4 | 15 | 50 | 62 | 229 | 24 | 391 |
| **0.9-1.0** | 0 | 0 | 0 | 0 | 1 | 1 | 11 | 22 | 45 | 80 |
| ***Total*** | 28 | 12 | 28 | 73 | 416 | 489 | 304 | 270 | 77 | 1797 |

**Table S2D**. Number of patients with transitions from Markov-states to Markov-states, based on updates between the yearly assessments of treatment with a TNFi in combination with sDMARD. These transitions are taking place in three out of four quarters of the simulation.

| Markov-state previous | Markov-state in subsequent quarter | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Synthetic** | **No DMARD** | **0.3-0.4** | **0.4-0.5** | **0.5-0.6** | **0.6-0.7** | **0.7-0.8** | **0.8-0.9** | **0.9-1.0** | ***Total*** |
| **0.3-0.4** | 1 | 0 | 63 | 6 | 3 | 1 | 0 | 1 | 0 | 75 |
| **0.4-0.5** | 7 | 0 | 8 | 174 | 30 | 7 | 3 | 3 | 0 | 232 |
| **0.5-0.6** | 22 | 3 | 4 | 23 | 1219 | 62 | 18 | 8 | 0 | 1359 |
| **0.6-0.7** | 15 | 9 | 0 | 11 | 54 | 1481 | 43 | 21 | 2 | 1636 |
| **0.7-0.8** | 8 | 1 | 1 | 5 | 16 | 35 | 943 | 15 | 3 | 1027 |
| **0.8-0.9** | 6 | 2 | 0 | 2 | 3 | 20 | 18 | 1206 | 9 | 1266 |
| **0.9-1.0** | 0 | 3 | 0 | 1 | 0 | 0 | 1 | 6 | 257 | 268 |
| ***Total*** | 59 | 18 | 76 | 222 | 1325 | 1606 | 1026 | 1260 | 271 | 5863 |

**Supplementary table 3. Reasons for discontinuation from NOR-DMARD and assumed continuation in model**

| Reason for discontinuation | Number of patients (with SF-6D data) | Percent | Mean SF-6D at discontinuation | Assumed continuation in model |
|---|---|---|---|---|
| Lack of efficacy | 82 (67) | 29 | 0.57 | TNFi + sDMARD or other biologic |
| Side effect of treatment | 80 (69) | 28 | 0.62 | sDMARD |
| Combination of side- effect and lack of efficacy | 8 (6) | 3 | 0.55 | sDMARD |
| Patient´s desire | 22 (15) | 8 | 0.69 | No DMARD treatment for 1 year |
| Remission of disease | 8 (7) | 3 | 0.70 | No DMARD treatment for 1 year |
| Other | 7 (4) | 2 | 0.59 | TNFi + sDMARD or other biologic |
| Unknown | 3 (3) | 1 | 0.63 | sDMARD |
| Death | 6 (4) | 2 | | |
| Medical event | 17 (13) | 6 | 0.57 | sDMARD |
| Drop-out | 40 (22) | 14 | 0.66 | TNFi + sDMARD or other biologic |
| (Wish of) pregnancy | 12 (11) | 4 | 0.75 | No DMARD treatment for 1 year |
| Total | 285 (221) | 100 | 0.62 | |

Source: NOR-DMARD

**Supplementary table 4. Overview of treatment types and number of RA-patients who started treatment with TNFi plus sDMARDs at start and after 1, 5 and 10 years of follow-up in NOR-DMARD.**

| Treatment type | Number starting treatment | After 1 year | After 5 years | At 10 years |
|---|---|---|---|---|
| TNFi + sDMARD | 860 | 591 | 172 | 11 |
| TNFi mono | 0 | 11 | 8 | 1 |
| Biologic | 0 | 20 | 25 | 1 |
| Synthetic | 0 | 17 | 8 | 0 |
| **Total DMARD** | **860** | **639** | **213** | **13** |
| Total with SF-6D | **810** | **508** | **178** | **11** |
| Total with EQ-5D | **456** | **256** | **64** | **2** |

Source: NOR-DMARD

**Supplementary table 5. Treatment types in RA patients who started treatment with TNF inhibitor plus synthetic DMARDs at baseline and after 1, 7 and 9 years of treatment. Inclusion criteria: SF-6D at start and at follow-up. Data from NOR-DMARD.**

| Treatment type | Start of treatment n (%) | After 1 year n (%) | After 7 years n (%) | After 9 years n (%) |
|---|---|---|---|---|
| Etanercept | | 6 (1) | 3 (3) | 1 (5) |
| Gold thiomalate | | 1 (0) | | |
| Leflunomide | | 1 (0) | | |
| Methotrexate | | 8 (2) | 4 (4) | |
| Sulfasalazine | | 1 (0) | | |
| MTX + anti-malarial | | 1 (0) | | |
| MTX + etanercept | 298 (37) | 179 (35) | 30 (31) | 13 (59) |
| MTX + infliximab | 211 (26) | 133 (26) | 19 (20) | 2 (9) |
| Adalimumab | | 2 (0) | 1 (1) | 1 (5) |
| MTX + adalimumab | 185 (23) | 116 (23) | 21 (22) | 2 (9) |
| Rituximab | | 2 (0) | 2 (2) | 1 (5) |
| Rituximab + MTX | | 10 (2) | 8 (8) | 2 (9) |
| Dmard combinations without MTX | | 1 (0) | | |
| Etanercept + leflunomide | 15 (2) | 5 (1) | | |
| Other combinations with rituximab | | | 1 (1) | |
| Abatacept+ MTX | | 1 (0) | 1 (1) | |
| Etanercept + sulfasalazine | 17 (2) | 3 (1) | 2 (2) | |
| Etanercept + azathioprine | 1 (0) | | | |
| Etanercept + hydroxychloquine | 2 (0) | 3 (1) | | |
| Etanercept + multiple DMARDs | 1 (0) | 1 (0) | | |
| Infliximab + leflunomide | 2 (0) | 1 (0) | | |
| Infliximab + sulfasalazine | 4 (0) | 2 (0) | 1 (1) | |
| Infliximab + azathioprine | 1 (0) | 2 (0) | | |
| Infliximab + multiple DMARDs | 3 (0) | 2 (0) | | |
| Adalimumab + leflunomide | 7 (1) | 4 (1) | | |
| Adalimumab + sulfasalazine | 8 (1) | 4 (1) | | |
| Adalimimab + azathioprine | 3 (0) | 1 (0) | | |
| Adalimumab + hydroxychloquine | 1 (0) | | | |
| Adalimumab + multiple DMARDs | 2 (0) | 1 (0) | | |
| Infliximab + gold thiomalate | | 1 (0) | | |
| Etanercept + gold thiomalate | 1 (0) | | | |
| Tocilizumab + MTX | | 1 (0) | 1 (1) | |
| Golimumab + MTX | 31 (4) | 12 (2) | | |
| Golimumab + leflunomide | 2 (0) | | | |
| Certolizumab + MTX | 13 (2) | 3 (1) | 2 (2) | |
| Certolizumab + leflunomide | 1 (0) | | | |
| Certolizumab + other DMARDs | 1 (0) | | | |
| Total | 810 | 508 | 96 | 22 |

**Supplementary table 6. Costs per unit in 2012 prices (€ 1.00=NOK 7.43)**

| Type of cost | Unit | Cost per unit in NOK(€)* | Information source |
|---|---|---|---|
| *Health care costs* | | | |
| Cost of infusion, exclusive drug cost | Visit | 1,629 (219) | The Norwegian Directorate of Health – telephone conversation. |
| General practitioner visit | Visit | 424 (57) | The Norwegian Medical Association (1) |
| Hospital stay | Stay** | 39,928 (5,374) | The Norwegian Directorate of Health(2) |
| Imaging examinations | Examination | 808 (109) | The Norwegian Directorate of Health-Regulations for financing outpatient radiology 2012(3) |
| Outpatient visit | Visit** | 1,262 (170) | The Norwegian Directorate of Health(2) |
| Private rheumatologist, 1st visit, comprehensive examination | Visit | 1,247 (168) | The Norwegian Medical Association(1) and the South-Eastern Norway Regional Health Authority |
| Private rheumatologist, 2nd and following visits | Visit | 986 (133) | The Norwegian Medical Association(1) and the South-Eastern Norway Regional Health Authority |
| Physiotherapy, 1st visit, comprehensive examination | Visit | 560 (75) | The Norwegian Physiotherapist Association |
| Physiotherapy, 2nd and following visits | Visit | 304 (41) | The Norwegian Physiotherapist Association |
| Rehabilitation stay, 21 days | Stay** | 45,248 (6,090) | Skogli Rehabilitation Centre |
| Travel cost (in addition to all other costs, except for imaging examinations) | Travel for one visit | 489(66) | The Corporate Health Centre for Patient Travel |
| *Production losses* | | | |
| Yearly income | One year | 470,900 (63,378) | Statistics Norway |

*Inclusive patient co-payments and exclusive travel expenses.
**The cost was calculated by the weight according to the activity-based funding.

**Supplementary table 7. Cost components in health states used in the model shown in Fig. 1.**

**Costs per three month period**
**Strategy:**
**TNFi+sDMARD**

1 €=7.43 NOK

| **sDMARD** | **NOK** | | | | **€** | | | |
|---|---|---|---|---|---|---|---|---|
| **Cost item** | **Mean** | **Std. Err.** | **[95% Conf.** | **Interval]** | **Mean** | **Std. Err.** | **[95% Conf.** | **Interval]** |
| Direct | 12 635 | 1 398 | 9 880 | 15 391 | 1 701 | 188 | 1 330 | 2 071 |
| Dmard | 775 | 109 | 561 | 989 | 104 | 15 | 76 | 133 |
| GP visit | 1 460 | 183 | 1 099 | 1 821 | 196 | 25 | 148 | 245 |
| Hospital stay (general) | 2 913 | 615 | 1 702 | 4 124 | 392 | 83 | 229 | 555 |
| Outpatient visit (general) | 434 | 80 | 276 | 592 | 58 | 11 | 37 | 80 |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 251 | 57 | 139 | 363 | 34 | 8 | 19 | 49 |
| Physiotherapy | 1 808 | 349 | 1 120 | 2 496 | 243 | 47 | 151 | 336 |
| Production loss | 56 275 | 4 968 | 46 484 | 66 065 | 7 574 | 669 | 6 256 | 8 892 |
| Imaging | 182 | 24 | 134 | 230 | 24 | 3 | 18 | 31 |
| Rehabilitation | 618 | 355 | -82 | 1 318 | 83 | 48 | -11 | 177 |
| Outpatient visit (rheumatology) | 371 | 67 | 239 | 502 | 50 | 9 | 32 | 68 |
| Hospital stay (rheumatology) | 3 823 | 903 | 2 043 | 5 604 | 515 | 122 | 275 | 754 |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 68 910 | 5 456 | 58 158 | 79 662 | 9 275 | 734 | 7 827 | 10 722 |
| **No DMARD** | | | | | | | | |
| Direct | 15 357 | 5 769 | 3 577 | 27 137 | 2 067 | 777 | 481 | 3 652 |
| Dmard | 0 | 0 | . | . | 0 | 0 | . | . |
| GP visit | 616 | 175 | 258 | 973 | 83 | 24 | 35 | 131 |
| Hospital stay (general) | 6 124 | 3 013 | -28 | 12 275 | 824 | 405 | -4 | 1 652 |
| Outpatient visit (general) | 849 | 340 | 155 | 1 543 | 114 | 46 | 21 | 208 |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 125 | 24 | 76 | 174 | 17 | 3 | 10 | 23 |
| Physiotherapy | 938 | 460 | -2 | 1 878 | 126 | 62 | 0 | 253 |
| Production loss | 89 150 | 13 906 | 60 757 | 117 543 | 11 999 | 1 872 | 8 177 | 15 820 |
| Imaging | 410 | 237 | -74 | 894 | 55 | 32 | -10 | 120 |
| Rehabilitation | 2 079 | 2 079 | -2 166 | 6 324 | 280 | 280 | -291 | 851 |
| Outpatient visit (rheumatology) | 849 | 292 | 253 | 1 445 | 114 | 39 | 34 | 195 |
| Hospital stay (rheumatology) | 3 368 | 2 589 | -1 918 | 8 654 | 453 | 348 | -258 | 1 165 |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 104 507 | 16 143 | 71 546 | 137 469 | 14 066 | 2 173 | 9 629 | 18 502 |

**TNFi+sDMARD - Level ≥ 0.296 and ≤0.4**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 37 817 | 3 147 | 31 601 | 44 034 | 5 090 | 424 | 4 253 | 5 926 |
| Dmard | 17 745 | 903 | 15 960 | 19 529 | 2 388 | 122 | 2 148 | 2 628 |
| GP visit | 1 325 | 149 | 1 030 | 1 619 | 178 | 20 | 139 | 218 |
| Hospital stay (general) | 1 023 | 330 | 372 | 1 674 | 138 | 44 | 50 | 225 |
| Outpatient visit (general) | 377 | 95 | 190 | 564 | 51 | 13 | 26 | 76 |
| Infusion | 496 | 140 | 220 | 772 | 67 | 19 | 30 | 104 |
| Other drugs | 334 | 62 | 211 | 458 | 45 | 8 | 28 | 62 |
| Physiotherapy | 1 888 | 362 | 1 174 | 2 603 | 254 | 49 | 158 | 350 |
| Production loss | 114 025 | 6 025 | 102 122 | 125 927 | 15 347 | 811 | 13 745 | 16 949 |
| Imaging | 250 | 35 | 182 | 319 | 34 | 5 | 24 | 43 |
| Rehabilitation | 1 230 | 519 | 204 | 2 256 | 166 | 70 | 28 | 304 |
| Outpatient visit (rheumatology) | 912 | 149 | 618 | 1 205 | 123 | 20 | 83 | 162 |
| Hospital stay (rheumatology) | 12 023 | 2 661 | 6 767 | 17 279 | 1 618 | 358 | 911 | 2 326 |
| Private rheumatologist | 22 | 15 | -9 | 53 | 3 | 2 | -1 | 7 |
| Total | 151 842 | 7 414 | 137 197 | 166 487 | 20 436 | 998 | 18 465 | 22 407 |

**TNFi+sDMARD - Level >0.4 and ≤0.5**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 48 363 | 2 223 | 43 992 | 52 733 | 6 509 | 299 | 5 921 | 7 097 |
| Dmard | 20 502 | 448 | 19 622 | 21 383 | 2 759 | 60 | 2 641 | 2 878 |
| GP visit | 1 099 | 82 | 938 | 1 259 | 148 | 11 | 126 | 169 |
| Hospital stay (general) | 2 078 | 483 | 1 128 | 3 027 | 280 | 65 | 152 | 407 |
| Outpatient visit (general) | 229 | 41 | 149 | 309 | 31 | 5 | 20 | 42 |
| Infusion | 455 | 78 | 302 | 608 | 61 | 10 | 41 | 82 |
| Other drugs | 247 | 12 | 223 | 270 | 33 | 2 | 30 | 36 |
| Physiotherapy | 2 587 | 263 | 2 069 | 3 104 | 348 | 35 | 278 | 418 |
| Production loss | 114 735 | 3 545 | 107 768 | 121 703 | 15 442 | 477 | 14 504 | 16 380 |
| Imaging | 275 | 31 | 215 | 336 | 37 | 4 | 29 | 45 |
| Rehabilitation | 1 389 | 309 | 781 | 1 997 | 187 | 42 | 105 | 269 |
| Outpatient visit (rheumatology) | 749 | 79 | 593 | 905 | 101 | 11 | 80 | 122 |
| Hospital stay (rheumatology) | 18 603 | 1 850 | 14 966 | 22 240 | 2 504 | 249 | 2 014 | 2 993 |
| Private rheumatologist | 8 | 6 | -3 | 19 | 1 | 1 | 0 | 3 |
| Total | 163 098 | 4 284 | 154 677 | 171 518 | 21 951 | 577 | 20 818 | 23 085 |

**TNFi+sDMARD - Level >0.5 and ≤0.6**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 36 883 | 733 | 35 446 | 38 320 | 4 964 | 99 | 4 771 | 5 157 |
| Dmard | 19 542 | 220 | 19 110 | 19 973 | 2 630 | 30 | 2 572 | 2 688 |
| GP visit | 761 | 26 | 711 | 811 | 102 | 3 | 96 | 109 |
| Hospital stay (general) | 2 076 | 189 | 1 705 | 2 447 | 279 | 25 | 229 | 329 |
| Outpatient visit (general) | 317 | 19 | 279 | 355 | 43 | 3 | 37 | 48 |
| Infusion | 393 | 29 | 336 | 449 | 53 | 4 | 45 | 60 |
| Other drugs | 274 | 12 | 251 | 297 | 37 | 2 | 34 | 40 |
| Physiotherapy | 2 056 | 97 | 1 865 | 2 247 | 277 | 13 | 251 | 302 |
| Production loss | 109 108 | 1 577 | 106 016 | 112 199 | 14 685 | 212 | 14 269 | 15 101 |
| Imaging | 246 | 14 | 219 | 274 | 33 | 2 | 29 | 37 |
| Rehabilitation | 1 480 | 140 | 1 205 | 1 754 | 199 | 19 | 162 | 236 |
| Outpatient visit (rheumatology) | 688 | 29 | 630 | 745 | 93 | 4 | 85 | 100 |
| Hospital stay (rheumatology) | 8 966 | 563 | 7 863 | 10 069 | 1 207 | 76 | 1 058 | 1 355 |
| Private rheumatologist | 3 | 2 | 0 | 6 | 0 | 0 | 0 | 1 |
| Total | 145 990 | 1 789 | 142 482 | 149 499 | 19 649 | 241 | 19 177 | 20 121 |

**TNFi+sDMARD - Level >0.6 and ≤0.7**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 34 315 | 599 | 33 141 | 35 489 | 4 618 | 81 | 4 460 | 4 776 |
| Dmard | 19 588 | 217 | 19 162 | 20 015 | 2 636 | 29 | 2 579 | 2 694 |
| GP visit | 671 | 20 | 633 | 710 | 90 | 3 | 85 | 96 |
| Hospital stay (general) | 2 020 | 150 | 1 726 | 2 314 | 272 | 20 | 232 | 311 |
| Outpatient visit (general) | 373 | 20 | 334 | 412 | 50 | 3 | 45 | 55 |
| Infusion | 294 | 22 | 250 | 338 | 40 | 3 | 34 | 45 |
| Other drugs | 209 | 10 | 190 | 229 | 28 | 1 | 26 | 31 |
| Physiotherapy | 2 067 | 93 | 1 886 | 2 249 | 278 | 12 | 254 | 303 |
| Production loss | 88 941 | 1 522 | 85 956 | 91 927 | 11 971 | 205 | 11 569 | 12 372 |
| Imaging | 188 | 9 | 170 | 206 | 25 | 1 | 23 | 28 |
| Rehabilitation | 1 109 | 127 | 860 | 1 359 | 149 | 17 | 116 | 183 |
| Outpatient visit (rheumatology) | 604 | 26 | 552 | 655 | 81 | 4 | 74 | 88 |
| Hospital stay (rheumatology) | 7 139 | 445 | 6 267 | 8 012 | 961 | 60 | 843 | 1 078 |
| Private rheumatologist | 9 | 6 | -2 | 19 | 1 | 1 | 0 | 3 |
| Total | 123 256 | 1 654 | 120 014 | 126 499 | 16 589 | 223 | 16 153 | 17 025 |

**TNFi+sDMARD - Level >0.7 and ≤0.8**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 31 563 | 735 | 30 122 | 33 005 | 4 248 | 99 | 4 054 | 4 442 |
| Dmard | 19 441 | 210 | 19 028 | 19 853 | 2 617 | 28 | 2 561 | 2 672 |
| GP visit | 570 | 23 | 524 | 616 | 77 | 3 | 71 | 83 |
| Hospital stay (general) | 1 760 | 238 | 1 293 | 2 226 | 237 | 32 | 174 | 300 |
| Outpatient visit (general) | 225 | 19 | 189 | 261 | 30 | 2 | 25 | 35 |
| Infusion | 282 | 30 | 222 | 341 | 38 | 4 | 30 | 46 |
| Other drugs | 183 | 7 | 168 | 197 | 25 | 1 | 23 | 27 |
| Physiotherapy | 1 762 | 108 | 1 551 | 1 974 | 237 | 14 | 209 | 266 |
| Production loss | 81 481 | 1 942 | 77 672 | 85 290 | 10 967 | 261 | 10 454 | 11 479 |
| Imaging | 196 | 21 | 154 | 237 | 26 | 3 | 21 | 32 |
| Rehabilitation | 624 | 97 | 433 | 815 | 84 | 13 | 58 | 110 |
| Outpatient visit (rheumatology) | 480 | 30 | 421 | 539 | 65 | 4 | 57 | 73 |
| Hospital stay (rheumatology) | 6 007 | 579 | 4 871 | 7 144 | 809 | 78 | 656 | 961 |
| Private rheumatologist | 5 | 2 | 0 | 9 | 1 | 0 | 0 | 1 |
| Total | 113 044 | 2 070 | 108 984 | 117 105 | 15 215 | 279 | 14 668 | 15 761 |

**TNFi+sDMARD - Level >0.8 and ≤0.9**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 30 371 | 1 159 | 28 098 | 32 643 | 4 088 | 156 | 3 782 | 4 393 |
| Dmard | 19 545 | 203 | 19 142 | 19 947 | 2 630 | 27 | 2 576 | 2 685 |
| GP visit | 498 | 22 | 455 | 541 | 67 | 3 | 61 | 73 |
| Hospital stay (general) | 1 041 | 124 | 798 | 1 284 | 140 | 17 | 107 | 173 |
| Outpatient visit (general) | 250 | 20 | 211 | 289 | 34 | 3 | 28 | 39 |
| Infusion | 204 | 22 | 161 | 247 | 27 | 3 | 22 | 33 |
| Other drugs | 142 | 5 | 134 | 151 | 19 | 1 | 18 | 20 |
| Physiotherapy | 1 142 | 81 | 983 | 1 302 | 154 | 11 | 132 | 175 |
| Production loss | 70 400 | 1 756 | 66 956 | 73 844 | 9 475 | 236 | 9 012 | 9 939 |
| Imaging | 123 | 8 | 107 | 139 | 17 | 1 | 14 | 19 |
| Rehabilitation | 2 606 | 1 004 | 636 | 4 576 | 351 | 135 | 86 | 616 |
| Outpatient visit (rheumatology) | 417 | 22 | 374 | 460 | 56 | 3 | 50 | 62 |
| Hospital stay (rheumatology) | 4 385 | 483 | 3 437 | 5 333 | 590 | 65 | 463 | 718 |
| Private rheumatologist | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| Total | 100 771 | 2 105 | 96 643 | 104 898 | 13 563 | 283 | 13 007 | 14 118 |

**TNFi+sDMARD - Level >0.9 and ≤1.0**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 23 169 | 777 | 21 642 | 24 696 | 3 118 | 105 | 2 913 | 3 324 |
| Dmard | 18 635 | 242 | 18 159 | 19 112 | 2 508 | 33 | 2 444 | 2 572 |
| GP visit | 263 | 26 | 211 | 315 | 35 | 4 | 28 | 42 |
| Hospital stay (general) | 207 | 89 | 32 | 383 | 28 | 12 | 4 | 52 |
| Outpatient visit (general) | 126 | 22 | 82 | 170 | 17 | 3 | 11 | 23 |
| Infusion | 206 | 39 | 130 | 283 | 28 | 5 | 17 | 38 |
| Other drugs | 59 | 6 | 47 | 71 | 8 | 1 | 6 | 10 |
| Physiotherapy | 462 | 124 | 219 | 706 | 62 | 17 | 29 | 95 |
| Production loss | 31 737 | 2 787 | 26 257 | 37 218 | 4 272 | 375 | 3 534 | 5 009 |
| Imaging | 66 | 8 | 50 | 82 | 9 | 1 | 7 | 11 |
| Rehabilitation | 235 | 101 | 36 | 433 | 32 | 14 | 5 | 58 |
| Outpatient visit (rheumatology) | 210 | 31 | 148 | 272 | 28 | 4 | 20 | 37 |
| Hospital stay (rheumatology) | 2 694 | 708 | 1 303 | 4 086 | 363 | 95 | 175 | 550 |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 54 906 | 2 809 | 49 383 | 60 430 | 7 390 | 378 | 6 646 | 8 133 |

**Strategy: sDMARD**

| | **NOK** | | | | **€** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Mean** | **Std. Err.** | **[95% Conf.** | **Interval]** | **Mean** | **Std. Err.** | **[95% Conf.** | **Interval]** |
| **sDMARD - Level ≥ 0.296 and ≤0.4** | | | | | | | | |
| Direct | 18 819 | 9 182 | -968 | 38 606 | 2 533 | 1 236 | -130 | 5 196 |
| Dmard | 523 | 150 | 199 | 847 | 70 | 20 | 27 | 114 |
| GP visit | 913 | 449 | -54 | 1 880 | 123 | 60 | -7 | 253 |
| Hospital stay (general) | 5 052 | 5 052 | -5 835 | 15 939 | 680 | 680 | -785 | 2 145 |
| Outpatient visit (general) | 0 | 0 | . | . | 0 | 0 | . | . |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 272 | 49 | 166 | 378 | 37 | 7 | 22 | 51 |
| Physiotherapy | 1 205 | 1 205 | -1 392 | 3 802 | 162 | 162 | -187 | 512 |
| Production loss | 113 310 | 19 725 | 70 806 | 155 815 | 15 250 | 2 655 | 9 530 | 20 971 |
| Imaging | 202 | 117 | -49 | 453 | 27 | 16 | -7 | 61 |
| Rehabilitation | 0 | 0 | . | . | 0 | 0 | . | . |
| Outpatient visit (rheumatology) | 547 | 264 | -21 | 1 115 | 74 | 35 | -3 | 150 |
| Hospital stay (rheumatology) | 10 104 | 4 519 | 367 | 19 842 | 1 360 | 608 | 49 | 2 671 |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 132 129 | 20 485 | 87 986 | 176 272 | 17 783 | 2 757 | 11 842 | 23 724 |

**sDMARD - Level >0.4 and ≤0.5**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 21 711 | 5 325 | 10 853 | 32 570 | 2 922 | 717 | 1 461 | 4 384 |
| Dmard | 312 | 70 | 170 | 453 | 42 | 9 | 23 | 61 |
| GP visit | 1 557 | 482 | 574 | 2 541 | 210 | 65 | 77 | 342 |
| Hospital stay (general) | 1 189 | 1 189 | -1 235 | 3 613 | 160 | 160 | -166 | 486 |
| Outpatient visit (general) | 258 | 131 | -9 | 524 | 35 | 18 | -1 | 71 |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 222 | 44 | 133 | 312 | 30 | 6 | 18 | 42 |
| Physiotherapy | 2 624 | 847 | 896 | 4 352 | 353 | 114 | 121 | 586 |
| Production loss | 95 350 | 13 537 | 67 747 | 122 954 | 12 833 | 1 822 | 9 118 | 16 548 |
| Imaging | 618 | 174 | 263 | 972 | 83 | 23 | 35 | 131 |
| Rehabilitation | 1 345 | 1 345 | -1 398 | 4 088 | 181 | 181 | -188 | 550 |
| Outpatient visit (rheumatology) | 1 648 | 377 | 880 | 2 416 | 222 | 51 | . | . |
| Hospital stay (rheumatology) | 11 887 | 3 632 | 4 482 | 19 293 | 1 600 | 489 | 603 | 2 597 |
| Private rheumatologist | 51 | 51 | -53 | 155 | 7 | 7 | -7 | 21 |
| Total | 117 062 | 16 594 | 83 225 | 150 899 | 15 755 | 2 233 | 11 201 | 20 309 |

**sDMARD - Level >0.5 and ≤0.6**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 15 913 | 3 326 | 9 315 | 22 511 | 2 142 | 448 | 1 254 | 3 030 |
| Dmard | 575 | 217 | 144 | 1 007 | 77 | 29 | 19 | 135 |
| GP visit | 685 | 107 | 472 | 897 | 92 | 14 | 64 | 121 |
| Hospital stay (general) | 2 332 | 1 333 | -313 | 4 976 | 314 | 179 | -42 | 670 |
| Outpatient visit (general) | 455 | 190 | 78 | 831 | 61 | 26 | 10 | 112 |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 426 | 195 | 40 | 812 | 57 | 26 | 5 | 109 |
| Physiotherapy | 2 077 | 603 | 880 | 3 273 | 280 | 81 | 119 | 441 |
| Production loss | 99 047 | 7 420 | 84 329 | 113 766 | 13 331 | 999 | 11 350 | 15 312 |
| Imaging | 396 | 73 | 251 | 541 | 53 | 10 | . | . |
| Rehabilitation | 440 | 440 | -433 | 1 312 | 59 | 59 | -58 | 177 |
| Outpatient visit (rheumatology) | 1 515 | 236 | 1 048 | 1 983 | 204 | 32 | 141 | 267 |
| Hospital stay (rheumatology) | 6 995 | 1 948 | 3 131 | 10 860 | 941 | 262 | 421 | 1 462 |
| Private rheumatologist | 17 | 17 | -16 | 50 | 2 | 2 | -2 | 7 |
| Total | 114 960 | 8 037 | 99 018 | 130 903 | 15 472 | 1 082 | 13 327 | 17 618 |

**sDMARD - Level >0.6 and ≤0.7**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 12 345 | 2 400 | 7 552 | 17 138 | 1 662 | 323 | 1 016 | 2 307 |
| Dmard | 301 | 53 | 195 | 408 | 41 | 7 | 26 | 55 |
| GP visit | 712 | 135 | 442 | 981 | 96 | 18 | 59 | 132 |
| Hospital stay (general) | 1 783 | 1 321 | -855 | 4 421 | 240 | 178 | -115 | 595 |
| Outpatient visit (general) | 335 | 112 | 112 | 558 | 45 | 15 | 15 | 75 |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 180 | 27 | 125 | 234 | 24 | 4 | 17 | 32 |
| Physiotherapy | 784 | 311 | 163 | 1 404 | 105 | 42 | . | . |
| Production loss | 74 652 | 9 080 | 56 517 | 92 786 | 10 047 | 1 222 | 7 607 | 12 488 |
| Imaging | 451 | 71 | 310 | 592 | 61 | 9 | 42 | 80 |
| Rehabilitation | 1 345 | 944 | -540 | 3 231 | 181 | 127 | -73 | 435 |
| Outpatient visit (rheumatology) | 1 700 | 316 | 1 069 | 2 330 | 229 | 42 | 144 | 314 |
| Hospital stay (rheumatology) | 4 755 | 1 591 | 1 578 | 7 932 | 640 | 214 | 212 | 1 068 |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 86 996 | 9 042 | 68 938 | 105 055 | 11 709 | 1 217 | 9 278 | 14 139 |

**sDMARD - Level >0.7 and ≤0.8**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 12 450 | 3 915 | 4 433 | 20 466 | 1 676 | 527 | 597 | 2 754 |
| Dmard | 256 | 54 | 146 | 366 | 35 | 7 | 20 | 49 |
| GP visit | 412 | 126 | 155 | 670 | 55 | 17 | 21 | 90 |
| Hospital stay (general) | 1 304 | 1 304 | -1 366 | 3 974 | 175 | 175 | -184 | 535 |
| Outpatient visit (general) | 395 | 194 | -2 | 793 | 53 | 26 | 0 | 107 |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 198 | 54 | 88 | 308 | 27 | 7 | . | . |
| Physiotherapy | 93 | 67 | -45 | 231 | 13 | 9 | -6 | 31 |
| Production loss | 64 012 | 12 407 | 38 604 | 89 420 | 8 615 | 1 670 | 5 196 | 12 035 |
| Imaging | 495 | 128 | 233 | 757 | 67 | 17 | 31 | 102 |
| Rehabilitation | 2 951 | 2 051 | -1 250 | 7 152 | 397 | 276 | -168 | 963 |
| Outpatient visit (rheumatology) | 1 130 | 299 | 518 | 1 742 | 152 | 40 | 70 | 234 |
| Hospital stay (rheumatology) | 5 215 | 2 474 | 149 | 10 281 | 702 | 333 | 20 | 1 384 |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 76 462 | 13 356 | 49 111 | 103 812 | 10 291 | 1 798 | 6 610 | 13 972 |

**sDMARD - Level >0.8 and ≤0.9**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 8 333 | 2 736 | 2 663 | 14 002 | 1 121 | 368 | 358 | 1 885 |
| Dmard | 346 | 87 | 166 | 526 | 47 | 12 | 22 | 71 |
| GP visit | 438 | 224 | -27 | 903 | 59 | 30 | -4 | 122 |
| Hospital stay (general) | 0 | 0 | . | . | 0 | 0 | . | . |
| Outpatient visit (general) | 210 | 116 | -31 | 451 | 28 | 16 | . | . |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 213 | 60 | 89 | 336 | 29 | 8 | 12 | 45 |
| Physiotherapy | 1 743 | 1 075 | -484 | 3 970 | 235 | 145 | -65 | 534 |
| Production loss | 42 522 | 12 440 | 16 738 | 68 307 | 5 723 | 1 674 | 2 253 | 9 193 |
| Imaging | 678 | 354 | -56 | 1 412 | 91 | 48 | -7 | 190 |
| Rehabilitation | 0 | 0 | . | . | 0 | 0 | . | . |
| Outpatient visit (rheumatology) | 1 471 | 425 | 590 | 2 352 | 198 | 57 | 79 | 317 |
| Hospital stay (rheumatology) | 3 233 | 2 238 | -1 406 | 7 872 | 435 | 301 | -189 | 1 060 |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 50 855 | 13 091 | 23 721 | 77 988 | 6 845 | 1 762 | 3 193 | 10 496 |

**sDMARD - Level >0.9 and ≤1.0**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Direct | 6 146 | 1 262 | -201 463 | 213 754 | 827 | 170 | -27 115 | 28 769 |
| Dmard | 163 | 0 | . | . | 22 | 0 | . | . |
| GP visit | 1 598 | 1 141 | -186 196 | 189 391 | 215 | 154 | . | . |
| Hospital stay (general) | 0 | 0 | . | . | 0 | 0 | . | . |
| Outpatient visit (general) | 0 | 0 | . | . | 0 | 0 | . | . |
| Infusion | 0 | 0 | . | . | 0 | 0 | . | . |
| Other drugs | 223 | 9 | -1 227 | 1 672 | 30 | 1 | -165 | 225 |
| Physiotherapy | 1 317 | 1 317 | -215 435 | 218 069 | 177 | 177 | -28 995 | 29 350 |
| Production loss | 123 611 | 41 204 | -6 656 782 | 6 904 004 | 16 637 | 5 546 | -895 933 | 929 206 |
| Imaging | 0 | 0 | . | . | 0 | 0 | . | . |
| Rehabilitation | 0 | 0 | . | . | 0 | 0 | . | . |
| Outpatient visit (rheumatology) | 2 846 | 1 094 | -177 254 | 182 945 | 383 | 147 | -23 856 | 24 622 |
| Hospital stay (rheumatology) | 0 | 0 | . | . | 0 | 0 | . | . |
| Private rheumatologist | 0 | 0 | . | . | 0 | 0 | . | . |
| Total | 129 757 | 42 465 | -6 858 244 | 7 117 758 | 17 464 | 5 715 | -923 048 | 957 976 |

**Supplementary table 8. Parameters and distributions in the probabilistic sensitivity analysis**

| Parameter | Distribution | Parameter value (Costs in €)<br>Mean<br>Standard Error (SE) | Number of observations |
|---|---|---|---|
| Probability of changing health state from 1st to 2nd quarter | Dirichlet | Table S1a | 632 |
| Probability of changing health state from 2nd to 3rd quarter | Dirichlet | Table S1b | 506 |
| Probability of changing health state in yearly assessments after the 3rd quarter | Dirichlet | Table S1c | 1,797 |
| Probability of changing health state in between assessments after the 3rd quarter | Dirichlet | Table S1d | 5,863 |
| Costs including production losses for TNFi-patients who switch to sDMARDs | Gamma | 9,324.37<br>725.14 | 228 |
| Costs including production losses for TNFi-patients who switch to No DMARD | Gamma | 11,898.35<br>1,955.24 | 41 |
| Costs including production losses for TNFi-patients in state 0.3-0.4 | Gamma | 20,203.49<br>977.13 | 165 |
| Costs including production losses for TNFi-patients in state 0.4-0.5 | Gamma | 21,581.67<br>568.27 | 447 |
| Costs including production losses for TNFi-patients in state 0.5-0.6 | Gamma | 19,517.81<br>236.78 | 2283 |
| Costs including production losses for TNFi-patients in state 0.6-0.7 | Gamma | 16,577.51<br>218.72 | 2628 |
| Costs including production losses for TNFi-patients in state 0.7-0.8 | Gamma | 15,164.51<br>273.57 | 1593 |
| Costs including production losses for TNFi-patients in state 0.8-0.9 | Gamma | 13,474.78<br>280.25 | 1857 |
| Costs including production losses for TNFi-patients in state 0.9-1.0 | Gamma | 7,390.21<br>373.22 | 405 |
| Health care costs for TNFi-patients who switch to sDMARDs | Gamma | 1,657.83<br>184.01 | 228 |
| Health care costs for TNFi-patients who switch to No DMARD | Gamma | 1,659.40<br>621.09 | 42 |
| Health care costs for TNFi-patients in state 0.3-0.4 | Gamma | 4,994.49<br>406.47 | 166 |
| Health care costs for TNFi-patients in state 0.4-0.5 | Gamma | 6,292.71<br>286.19 | 453 |
| Health care costs for TNFi-patients in state 0.5-0.6 | Gamma | 4,893.70<br>95.66 | 2307 |
| Health care costs for TNFi-patients in state 0.6-0.7 | Gamma | 4,550.75<br>78.62 | 2640 |
| Health care costs for TNFi-patients in state 0.7-0.8 | Gamma | 4,200.31<br>95.98 | 1599 |
| Health care costs for TNFi-patients in state 0.8-0.9 | Gamma | 4,058.97<br>151.13 | 1889 |
| Health care costs for TNFi-patients in state 0.9-1.0 | Gamma | 3,165.28<br>110.19 | 411 |

| | | | |
|---|---|---|---|
| Costs including production losses for sDMARD patients in state 0.3-0.4 | Gamma | 17,783.20<br>2,757.08 | 16 |
| Costs including production losses for sDMARD patients in state 0.4-0.5 | Gamma | 15,755.28<br>2,233.43 | 34 |
| Costs including production losses for sDMARD patients in state 0.5-0.6 | Gamma | 15,472.45<br>1,081.64 | 104 |
| Costs including production losses for sDMARD patients in state 0.6-0.7 | Gamma | 11,708.81<br>1,216.99 | 68 |
| Costs including production losses for sDMARD patients in state 0.7-0.8 | Gamma | 10,290.92<br>1,797.56 | 31 |
| Costs including production losses for sDMARD patients in state 0.8-0.9 | Gamma | 6,844.52<br>1,761.91 | 25 |
| Costs including production losses for sDMARD patients in state 0.9-1.0 | Gamma | 11,685.30<br>6,654.42 | 3 |
| Health care costs for sDMARD patients in state 0.3-0.4 | Gamma | 2,532.83<br>1,235.86 | 16 |
| Health care costs for sDMARD patients in state 0.4-0.5 | Gamma | 2,839.26<br>700.87 | 35 |
| Health care costs for sDMARD patients in state 0.5-0.6 | Gamma | 2,168.25<br>438.27 | 107 |
| Health care costs for sDMARD patients in state 0.6-0.7 | Gamma | 1,637.66<br>319.18 | 69 |
| Health care costs for sDMARD patients in state 0.7-0.8 | Gamma | 1,801.25<br>525.38 | 32 |
| Health care costs for sDMARD patients in state 0.8-0.9 | Gamma | 1,121.47<br>368.17 | 25 |
| Health care costs for sDMARD patients in state 0.9-1.0 | Gamma | 594.11<br>252.83 | 3 |
| SF-6D utility in state 0.3-0.4 | Gamma | .3741372<br>.0038512 | 51 |
| EQ-5D utility in state 0.3-0.4 | Gamma | -.0306538<br>.0212116 | 26 |
| SF-6D utility in state 0.4-0.5 | Gamma | .4594679<br>.0021211 | 109 |
| EQ-5D utility in state 0.4-0.5 | Gamma | .2297925<br>.0394939 | 53 |
| SF-6D utility in state 0.5-0.6 | Gamma | .5550128<br>.0013444 | 312 |
| EQ-5D utility in state 0.5-0.6 | Gamma | .4529091<br>.0199665 | 165 |
| SF-6D utility in state 0.6-0.7 | Gamma | .6390674<br>.0021241 | 193 |
| EQ-5D utility in state 0.6-0.7 | Gamma | .5604194<br>.0227441 | 124 |
| SF-6D utility in state 0.7-0.8 | Gamma | .7407857<br>.0026047 | 84 |
| EQ-5D utility in state 0.7-0.8 | Gamma | .7451304<br>.0123842 | 46 |
| SF-6D utility in state 0.8-0.9 | Gamma | .8530877<br>.0044369 | 57 |

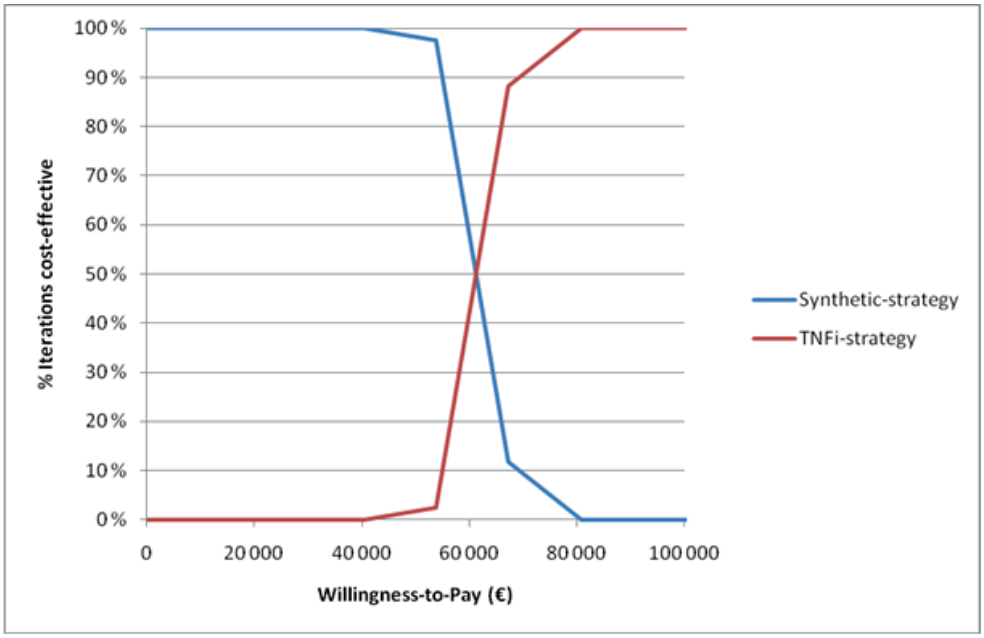
| | | | |
|---|---|---|---|
| EQ-5D utility in state 0.8-0.9 | Gamma | .7938718<br>.0093293 | 39 |
| SF-6D utility in state 0.9-1.0 | Gamma | .9395<br>.0049075 | 4 |
| EQ-5D utility in state 0.9-1.0 | Gamma | .841<br>.0819573 | 3 |
| SF-6D utility in state Synthetic DMARD | Gamma | .6352444<br>.0073723 | 225 |
| EQ-5D utility in state Synthetic DMARD | Gamma | .5530572<br>.0243754 | 105 |
| SF-6D utility in state No DMARD | Gamma | .7182188<br>.0170004 | 64 |
| EQ-5D utility in state No DMARD | Gamma | .7452759<br>.0603896 | 29 |

**Supplementary table 9. Cost-effectiveness acceptability**

| Type of cost | Type of utility instru-ment | Strategy | ICER € | Probability of being cost-effective at a WTP of €53,800 (NOK400,000) | WTP of €67,300 (NOK500,000) | WTP of €80,800 (NOK600,000) | WTP of €94,200 (NOK700,000) |
|---|---|---|---|---|---|---|---|
| Health care | SF-6D | Synthetic | | 1 | 1 | 1 | 0.256 |
| | | TNFi | 92,557 | 0 | 0 | 0 | 0.744 |
| | EQ-5D | Synthetic | | 0.971 | 0.097 | 0 | 0 |
| | | TNFi | 61,285 | 0.029 | 0.903 | 1 | 1 |
| Health care + produc-tion losses | SF-6D | Synthetic | | 0.890 | 0.110 | 0.001 | 0 |
| | | TNFi | 60,227 | 0.110 | 0.890 | 0.999 | 1 |
| | EQ-5D | Synthetic | | 0.001 | 0 | 0 | 0 |
| | | TNFi | 39,872 | 0.999 | 1 | 1 | 1 |

ICER: Incremental cost-effectiveness ratio

**Supplementary figure 1. Cost effectiveness acceptability curve. Health care costs in € and EQ-5D utilities. The probability for the TNFi-/synthetic-strategy to be cost-effective assuming different WTP-thresholds are illustrated in the figure.**



**Supplementary text 1**

**Costs for the TNFi-strategy:**

A start cost was added to all patients in the TNFi-strategy at the start of treatment, including obligatory examinations before the initiation of a bDMARD therapy. The cost of taking a chest x-ray was added for all patients. Additionally, the start cost included the cost of blood analyses (Hepatitis B and C), a test for Tuberculosis; either the Mantoux or the QuantiFERON® test (50% were supposed to take the Mantoux and 50 % the QuantiFERON® test), with a half hour wage for a nurse for taking the Mantoux test. The nurse's wage estimate was based on a public health nurse in wage category 30 in 2012. The price of the Mantoux test was taken from the price list of vaccinations for 2012 from The Norwegian Institute of Public Health. The prices for the Hepatitis B and C and the QuantiFERON® tests were given from the Chief Medical Officer at the Department of Microbiology at Oslo University Hospital in 2012. The start cost that was added to all patients included the blood analyses and the tuberculosis tests and totaled NOK737 (€99).

The quarterly cost estimate was based on patients from NOR-DMARD who started treatment with a TNFi plus sDMARDs, had no previous use of biologic therapy and were diagnosed with RA. After the first quarter, the cost estimate included patients who continued on this regimen, patients who switched to mono TNFi treatment, to another TNFi or changed to another biologic treatment and did not change diagnosis during the follow-up. The number of patients in the health care cost estimate was at start 802(the corresponding number of patients in the utility estimate was 810), after 5 years 221(230) and after 10 years 10(11). The number of patients in the total cost estimate was at start 789(810), after 5 years 221(230) and after 10 years 10(11).

If patients switched to sDMARD treatment they were transferred to the state "sDMARD" and received the mean cost for the patients on sDMARD treatment in NOR-DMARD after the failure of a TNFi, corresponding to their utility state, i.e state 0.60-0.69.

**Cost items for the TNFi-strategy and for the synthetic-strategy.**

*Pharmaceutical costs*

For each patient, all drugs used between assessments were summarized. The use of DMARDs was calculated by multiplying doses expressed in milligram by frequency (daily, weekly or monthly). Infusions of bDMARD were registered by date and dose and these were summarized for each patient and period. Concomitant medications (including analgesics, anti-inflammatory products and glucocorticoids), were registered by start and stop dates and added for each patient and period. For concomitant medications, the defined daily dose (DDD) from WHO Collaborating Centre for Drug Statistics Methodology was used. If the patient used the medication as a regular regimen, the DDD was counted as one and if the patient used the medication "as needed", the DDD was multiplied by 0.25.

Other cost components

Travel expenses were included in all outpatient examinations-, and treatments as well as in hospital-, and rehabilitation stays. Travel costs were not included for imaging examinations, which were often performed at the same time as an outpatient visit. We used the mean cost for all patient travels in Norway in 2012 according to the Corporate Health Centre for Patient Travel.

*Private practicing rheumatologist*

The cost of a visit to a private practicing rheumatologist was separated into 1) the first visit and 2) the second and following visits. The first visit included: The consultation fee for specialists, a full examination fee, fees for blood analyses (CRP, ALAT and GT)(1) and the operation subsidy divided by number of consultations. The following visits included the same cost components except for the full examination fee.

*General practitioner*

The cost of a visit to a General practitioner (GP) included the consultation fee for GPs, an extra fee for specialists, fees for blood analyses (CRP, ALAT and GT), the basic yearly public pay to the GP per patient divided by four (assuming each patient visits the GP on average four times a year).(1)

*Outpatient visits*

Visits to the Rheumatology and other outpatient clinics had the same DRG-weight in the Norwegian system and were priced accordingly. The DRG-weight was multiplied by the Unit Reimbursement 2012.(2)

*Hospital stays*

Two DRG-weights were used for "Specific inflammatory joint and spinal diseases", with and without a secondary diagnosis or complicating disease. The mean of the two hospital costs was used as the estimate for one hospital stay.(2)

*Imaging examinations*

The weights for costs associated with imaging examinations were found in the new regulations on financing outpatient radiology from 2012 that is performed at public health institutions and by health care facilities that receive operating subsidies from regional health authorities. The tariff RG1 from Norwegian Classification of Radiological Procedures, NCRP was used. We assumed that the patient had x-rays of the chest, hands and feet. The reimbursement of the RG1 plus the patient deductible are supposed to be 40% of the total cost, hence the remaining 60% was added to calculate the total price of the examination.(3)

*Physical therapy*

The cost of a visit to a physiotherapist was separated into 1) the first visit and 2) the second and following visits. The first visit included the cost of an examination plus the cost of an additional 30 minutes and the operating subsidy divided by number of treatment hours per year. The second and following visits included costs for exercises or massage therapy, supervised training and the operating subsidy divided by treatment half hours.

*Rehabilitation*

The cost estimate for a rehabilitation stay was based on a mean stay for a patient with RA at one of the largest rehabilitation centres in Norway: Skogli Rehabilitation Centre. It was assumed that the stay lasted for 21 days and that 1/3 of the patients had individual treatment and 2/3 followed a programme in a group.

*Production losses*

For production losses, we used the mean income in Norway in 2012 (NOK 470,900) (Statistics Norway http://www.ssb.no/lonnansatt). For patients who were unable to work, including patients on disability pension, rehabilitation and sick leave, the mean income plus the social cost of labour (40% of the income) were added as a production loss. For patients who were able to work part time, this productivity cost was reduced in proportion to time worked. We adopted the human capital approach, assuming that the production loss persisted throughout the analysed time period.(4)

**Costs for the synthetic-strategy:**

No start costs were added for patients in this strategy since a possible x-ray examination was supposed to have taken place when patients started treatment with a sDMARD. The modeling started at the point when patients were assumed to have failed two sDMARDs.

The cost estimate for patients in the synthetic-strategy was based on patients from NOR-DMARD who stated treatment with a TNFi, had no previous use of biologic therapy and were diagnosed with RA, **before** they started the treatment with a TNFi. The cost estimate was the mean of these patients' costs in the last three months before the start with a TNFi. The number of patients in the health care cost estimate in the three months preceding the start with a TNFi was 287. The number of patients in the cost estimate including productivity losses in the three months preceding the start with a TNFi was 281.

REFERENCES

1. Den norske legeforening [The Norwegian Medical Association]. Normaltariff for privat spesialistpraksis 2011-2012. [The normal tariff for private specialist practice]. 2012.
2. Helsedirektoratet [The Norwegian Directorate of Health]. Insatsstyrt finansiering 2012. [Activity Based Financing]. Oslo: 2011.
3. Regulations for financing outpatient radiology 2012 - state health care. 2011.
4. Drummond MF, Sculpher MJ, Stoddart GL, O'Brien BJ, Torrance GW. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford: Oxford University Press; 2005.

# 10 ERRATA

## 10.1 Paper II

1. Table 2 “Unit costs in 2010 prices”
   The cost per unit for “Physiotherapy, second and following visits” is presented in NOK (282). The number should have been in € (35). Thus, 282 should be replaced by 35.

2. Table 2 “Unit costs in 2010 prices”
   The exchange rate used in the table is €1.00=NOK 8.15 from 2010. The analyses are presented with an updated exchange rate after a revision of the article in 2011. The updated exchange rate from 2011 was €1.00=NOK 7.81. The updated exchange rate should have been used in table 2, as in the rest of the paper.

3. Table 3 “Direct and indirect cost components in euros over 2 years according to the diagnosis and synthetic or biologic DMARD”.
   The presented total mean costs in table 3, estimated by the HCA and the FCA, are the observed mean values, calculated by 10,000 bootstrapped estimates. The observed bootstrapped mean values are 0.27% different from the arithmetic mean values of the four half year periods summed together. The total mean costs in the table are consequently 0.27% different from the total mean costs in the main text and the abstract. The values presented in table 3 should have been the same as in the text, *i.e.* the arithmetic mean values.
   For example: The total 2-year costs for RA-patients on biologic treatment in the text are €121,900(HC) and in the table the total 2-year costs are €122,233(HC).

4. In Results, page 1622.
   ’The changes in total costs from the first to the last 6-month periods were significant in all diagnoses for both synthetic and biologic treatment ($p$<0.05).’

   This sentence should have been:
   ’The changes in total costs from the first to the last 6-month periods were significant in all diagnoses for biologic treatment ($p$<0.05).’