Health Economic Analyses Using Real-World Data in the Field of Rheumatology

Celine van de Laar

Health Economic Analyses using Real-World Data in the Field of Rheumatology

Celine van de Laar

Health Economic Analyses using Real-World Data in the Field of Rheumatology

Gezondheidseconomische analyses met real-world data binnen de reumatologie

Thesis

to obtain the degree of doctor from the

Erasmus university Rotterdam

by command of the

rector magnificus

Prof. dr. ir. Jantine Schuit

and in accordance with the decision of the Doctorate Board.

The public defence shall be held on

Friday July 4, 2025 at 10:30 hrs

by

Celine Johanna van de Laar

Born in Almelo, the Netherlands

Doctoral Committee:

Promotor:	Prof.dr. C. A. Uyl-De Groot
Other members:	Prof.dr. A.E.R.C.H. Boonen
	Prof.dr. A.E.A.M. Weel-Koenders
	Prof.dr.ir. C.T.B. Ahaus
Copromotors:	Dr. A.H. Oude Voshaar
	Dr. P.M. Ten Klooster

Chapter 1

Chapter 1 – Introduction

Health Economic Analyses with Real-World Data in the Field of Rheumatology

an introduction to the thesis by

Celine J. van de Laar

Rheumatology

Rheumatology is the medical specialty that is focused on conditions of the musculoskeletal system, not caused by trauma or accidents. Many inflammatory diseases are driven by auto-immune disorders, rheumatology is a field uniquely positioned to address such systemic auto-immune disorders. The spectrum of rheumatic diseases is extensive and diverse, ranging from degenerative disorders to metabolic conditions and inflammatory diseases.

Rheumatoid Arthritis

There are a multitude of different rheumatic diseases, but Rheumatoid Arthritis (RA) is the most well-known. It affects approximately 0.5%-1% of the world's population, affects women twice as much as men and is the prototype of both inflammatory rheumatic diseases and systemic auto immune diseases [1,2].

The basic pathological process of RA lies in the immune system. Driven largely by unknown genetic predispositions, combined with external factors, the immune system reacts with cells of the patient's body, frequently accompanied but not necessarily, by the production of rheumatoid factors (RF) and/or anti-Cyclic Citrullinated Peptide (a-CCP). Most likely explained by the genetic predisposition, RA is clustered in families. RA may develop at all ages, but usually presents in adulthood. RA can be triggered by hormonal changes around pregnancy and menopause [3]. People who develop rheumatic signs and symptoms often fear the prognosis as they fear RA, due to its well-known severity. However, rheumatic diseases differ widely and can show a range of different symptoms and characteristics. The diagnostic process is not always straightforward, as most rheumatic diseases lack a single distinguishing feature [4]. Mapping a multi-dimensional phenotype, integrating clinical symptoms, laboratory data, patient history, imaging, and genetic or environmental factors is essential towards proper diagnosis. Fortunately, the treatment of RA has improved greatly over the past decades [1]. The introduction of new treatments, control strategies, and patient involvement have improved both biomedical outcomes as well as outcomes that matter for patients [5–7].

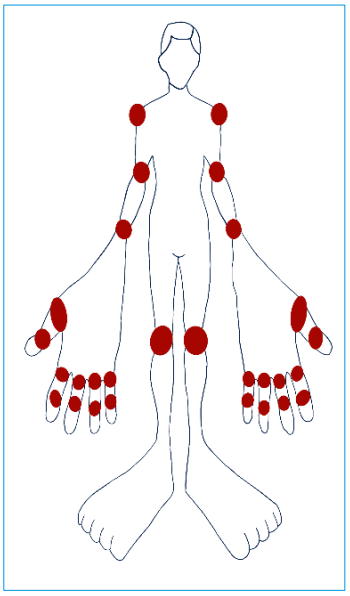


Figure 1 typical distribution of affected joints in Rheumatoid arthritis

Presumably as a result of autoimmunity, the chronic inflammatory process starts. Although many organs can be a target of the disease, typically the disease starts with inflamed joints. The disease is a chronic condition with almost permanently inflamed joints, however, not all joints are involved [1]. The distribution on affected joints is symmetrical and peripheral except for of the most distal joints of fingers and feet (Figure 1).

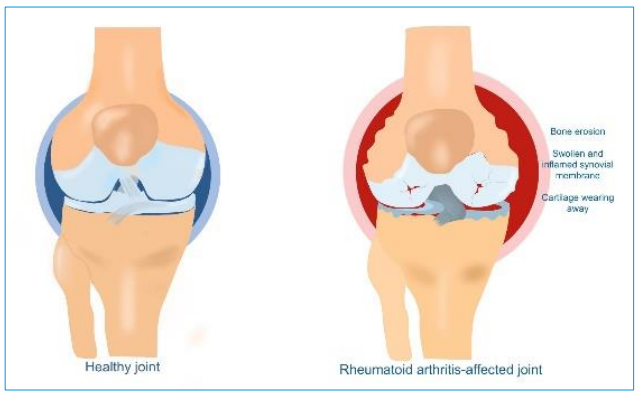


Figure 2 on the left a scheme of a healthy joint. On the right a joint with typical Rheumatoid arthritis features

Within the joints, the synovial membrane is the specific target of inflammation. During the inflammatory process, the synovium expands due to proliferation of blood vessels and inflammatory cells, which in turn leads to the increased production of synovial fluid of moderate quality (Figure 2). The local inflammation causes swelling, pain, elevated local temperature, and stiffness. The inflammation rarely becomes severe enough to the point that redness is visible, or fever is present. However, even at early disease stages, patients experience fatigue and impaired physical functionality due to the low grade inflammation, paired with pain.

Without treatment, the natural course of RA is variable. A small percentage of patients experience spontaneous drug-free remission after a first episode. The disease course in most patients is erratic and comes with remissions and exacerbations, alternating with calmer periods [8].

If untreated, the signs and symptoms of RA, caused by inflammation, lead to a temporary decline of functional capacity with regard to both work and daily life. The inflammation of synovial fluid can be destructive locally. In a considerable number of patients, cartilage, bone, capsules, and tendons are destroyed leading to long term damage and even permanent disability (Figure 2). Typically, eroding of the bony parts of the joint can be visualised and scored by plain x-rays (Figure 3) [9].

Up to the turn of the century, the treatment of RA was primarily symptomatic and the fate and prognosis of patients suffering from RA regarding their quality of life and work was disappointing. Presently, the good news is that various treatments, especially pharmacological, can relieve the pain and slow down – or even stop – the progression of the disease [8]. Therefore, an early diagnosis and well-controlled treatment is warranted and necessary. This will be further discussed hereafter.

Diagnosis

In order to prevent (long-term) damage, early diagnosis and adequate treatment is key. To ensure this, the general public and their attending general practitioners (GPs), need to be aware of the presenting signs and symptoms of RA. The key signs of joint inflammation: warmth, swelling, and pain can be accompanied by:

Stiffness of joints: especially in the morning after awakening, patients tend to only become flexible again after more than an hour, weakness: painful, stiff joints often end up not getting as much use, which can cause the muscles to get weaker over time, exhaustion: RA is a systemic inflammatory disease often causing fatigue and general physical weakness.

People confronted with these signs and symptoms should be alarmed and consult their GP. In turn, GPs should be trained and skilled in recognizing inflamed joints: arthritis. GPs are usually aware of a patient's family and smoking history, which are additional risk factors for rheumatic diseases. However, all patients suspected of arthritis should be referred to a rheumatologist by their GP, even those without additional risk factors. Additional investigations by a GP, like ordering laboratory tests for inflammation, RF or aCCP, or radiology before referral to a rheumatologist, have limited or no additional value at this point [10].

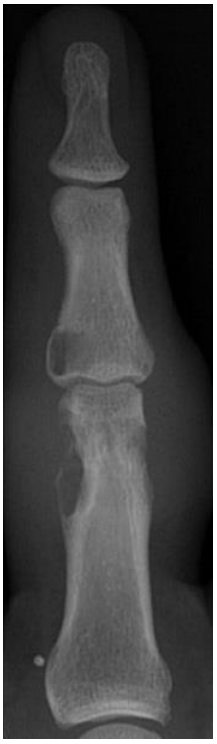


Figure 3 plain x-ray of peripheral joints showing typical erosions

The rheumatologist's primary task is to verify joint inflammation and then identify the various possible causes in order to make a clinical diagnosis of RA or the appropriate alternative explanation. After determining the diagnosis RA, the rheumatologist and allied health professionals inform the patient and their GP. The state of the disease and the prognosis will be assessed accordingly. Based upon the guidelines and in line with the needs, knowledge, skills, and preferences of the patient an individual treatment plan must be designed [11]. This process is referred to as shared decision making.

Treatment

The aim of this treatment plan is to reduce symptoms and to prevent unnecessary damage to improve outcomes that matter to patients. Since the execution of this treatment lies primarily in the hands of the patient, optimal self-efficacy needs a decision supported by the patient, which can be obtained by shared decision-making.

In order to alleviate the patient's discomfort, the treatment starts with pharmaceutical as well as non-pharmaceutical interventions. Non-steroidal anti-inflammatory drugs (NSAIDs) and corticosteroids have good short term effects and can be combined with physical- and ergotherapy. Effective symptomatic therapy not only relieves the patient burden almost immediately but also supports the trust in the rheumatologist and their team, which is a critical success factor towards reaching a long-term effective patient-physician relationship [10].

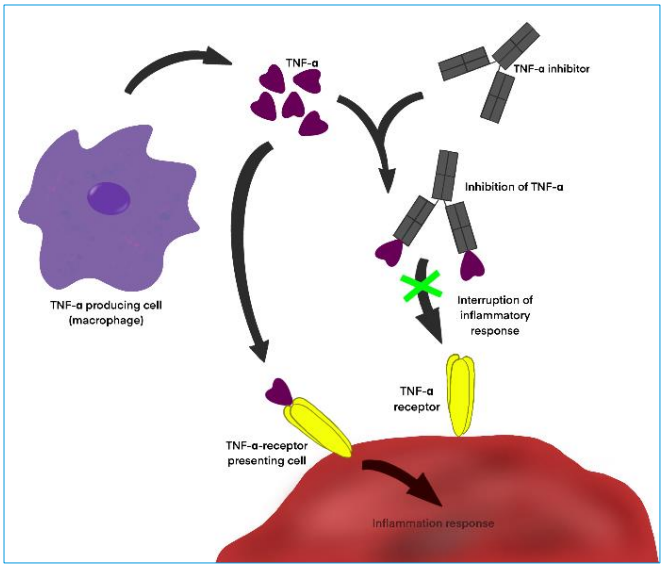


Figure 4: schematic overview of the working mechanism of TNFi (anti-TNF-alfa)

Guidelines also advise to immediately start disease-modifying treatment, alongside symptomatic treatment. Over the past two decades, multiple innovative Disease Modifying Anti-Rheumatic Drugs (DMARDs) have been developed and implemented. Especially the introduction of inhibitors of Tumor Necrosis Factor (TNFi) in the year 2000, changed the perspective of RA patients from *care* to a disease where *cure* is almost an option (see Figure 4). TNFi are classified as biological Disease Modifying Anti-Rheumatic Drugs (bDMARDs). Currently, various TNFi are available as originals but as of recently also as biosimilars. The introduction of biosimilars has significantly reduced the costs of TNFi treatment. Other pro-inflammatory targets in the complex system regulating immunity and inflammation have been identified and successfully targeted by bDMARDs as well [12]. However, all bDMARDs need to be administrated by injection and the majority of patients prefer oral medications and some patients even reject parenteral treatments. More recently, signalling pathways regulating immunity and inflammation have been identified and can be effectively targeted with oral drugs. These inhibitors of the Janus kinase (JAK) signal transducer and activator of transcription (JAK-STAT) pathway modulate the effects of multiple

cytokines and have been shown to have comparable or better results as compared with prototypes of bDMARDs that target, by inhibition, only one cytokine.

These JAK-inhibitors (JAKi), also called targeted synthetic DMARDs (tsDMARDs), are as of now only available as originators [13], however, the first patents and/or biosimilars are likely come to market in the near future. All drugs in the DMARD arsenal have been shown to inhibit not only disease activity but also long-term effects. The philosophical - mathematical principle that the integral over time of disease activity is related to outcomes has been demonstrated for different disease activity parameters, structural damage, and for outcomes that matter to patients. Moreover, in the Netherlands, classical life-threatening complications of RA, for example, amyloidosis and vasculitis, have been eradicated by the present effective treatments [14].

Apart from the pharmacological innovations described above, rheumatologists have learned how to apply these drugs in optimal strategies and how patients can be involved and motivated. Early diagnosis and treating to the target of remission (T2T) are the guiding principles, in addition to shared decision-making [15].

Gout

Introduction

A rheumatic disease with significantly different causes, symptoms, and treatment approach is gout. Contrasting RA, gout is an ancient disease. Only recently, a relevant increase in the medical and fundamental knowledge of the disease has made way for accurate diagnosis and innovations in treatment and prevention. Simply put, gout is a metabolic condition leading to a systemic inflammatory disease driven by the deposition of monosodium urate (MSU) crystals in tissues. Although hyperuricemia is the main pathogenic defect in gout, the majority of patients with hyperuricemia never develop gout. Formation of gout crystals consists of more than achieving a MSU concentration above a solubility concentration [16].

Diagnosis

Although gout crystals can be formed in all tissues, the joints and periarticular structure have a strong preference. A diagnosis made by microscopic identification of MSU crystals using polarized light is the gold standard, nonetheless the majority of patients obtain the diagnosis based upon the clinical picture of podagra. Gouty arthritis typically starts as an acute mono-arthritis of the first metatarsophalangeal (MTP-1) joint, referred to as podagra. The reddish coloration emphasizes the severity of the inflammation (see Figure 5). The symptoms of the first attacks

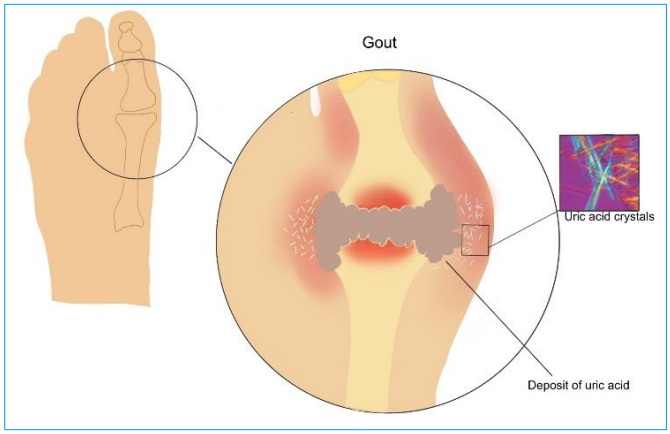


Figure 5 art impression of (peri)articular urate deposition in a podagra patient and an insert of visible mono sodium urate crystals as seen by polarised light microscopy.

resolve, even untreated, in a few days. After the first attack, recurrent attacks may reoccur in other and in more joints. Most flares of gouty arthritis are quickly relieved by NSAIDs or colchicine. But the dependency on pharmacotherapy will increase. When insufficiently, treated the deposition of MSU crystals continues and so called tophi are formed. MSU deposition in the kidney typically complicates this disease [17].

Gout, being a metabolic disease with systemic implications, is lifestyle related. Therefore, it is not surprising that the prevalence increases with increasing prosperity, obesity and metabolic syndrome. In combination with genetic predisposition, in specific populations, the prevalence of gout increases to 10%. Men are more frequently affected than women. The consumption of alcohol, use of diuretics and renal impairment are related to gout as well [18].

The formation of MSU crystals in the joint trigger gouty arthritis. These MSU crystals may activate inflammatory cells. These cells in turn produce cytokines driving inflammation, among others Interleukine-1 (IL-1). This results in increasing vascular permeability and vasodilatation, accelerating an acute intensive inflammation [19]. Resorption as well as encapsulation of crystals settles the acute attack, clinically experienced as remission [20]. Sustained hyperuricemia, increased MSU crystal depositions resulting in an increase in total body urate, is associated with chronic gout, characterized by chronic synovitis, bony erosions, cartilage damage, and tophi formation. Extra articular manifestations of chronic gout include cardiovascular and renal complications [21].

Globally, the gout diagnosis is made on clinical grounds only. Dutch rheumatologists and their national guidelines have adapted the more accurate but unique diagnosis mechanism using the gold standard of synovial fluid analysis by Polarized Light Microscopic diagnosis of crystals (Figure 5) [22].

Treatment

On the one hand, gout is the best-understood rheumatic disease. However, even though effective urate lowering therapies (ULT) are available, therapeutic failure is frequent. Treatment failures can be caused by an inappropriate diagnosis, inappropriate treatment strategy, or insufficient execution of the treatment strategy. Successful treatment selection and execution is driven by shared decision-making. Obviously, tolerability and adverse events in the gout population where comorbidities and comedications are prevalent call for safe and effective innovations, a currently unmet need [23].

Figure 6: Autumn Crocus (Colchicaneae) containing the active anti-inflammatory agent: Colchicine.

The prototype for symptomatic treatment of gout is an ancient drug called colchicine, historically extracted from the autumn crocus or colchicaneae (see Figure 6). Alternatively, NSAIDs and corticosteroids are used to control pain and inflammation.

Neither of these drugs are formally licensed for the treatment of gout flares and come with contraindications that are highly prevalent in the gout population. Only canakinumab, an inhibitor of IL-1 is formally licensed for the treatment of gout flares but is considered extremely expensive. Another, IL-1 inhibitor: anakinra, is proven to be effective for the treatment of gout but again, not licensed for this indication [24]. Hence, there is a pressing unmet need for effective, safe, licensed and affordable treatments.

Moreover, correction of hyperuricemia, the metabolic basis of gout, mechanistically should prevent recurrent attacks and complications of gout. To prevent de novo formation and dissolution of MSU crystals, Serum Urate Levels (SUL) should be lowered to values under the SUL saturation point. Both the American College of Rheumatology (ACR) and the European League against Rheumatism (EULAR) indicate that the SUL target is below 0.36 (mmol/l) in all gouty patients and below 0.30 (mmol/l) in severe gout patients. Like in RA, ULT in gout is a treatment to a target enabled by routine checks of SUL [25].

Urate lowering in gout is a long-term process and in order for it to be successful, it is dependent both on effective drugs as well as effective execution of the treatment, compliance and dosing. An educated patient who shares the decision for a specific ULT and everything that comes with that is key to a successful outcome. In addition to the pharmacotherapy, lifestyle changes are of major importance to improve the urate metabolism as well as preventing comorbidities that are lifestyle related [26].

Real World Data in Rheumatology

In the Netherlands, clinical scientists have strongly contributed to the innovations in rheumatology. prof. dr. Hans Valkenburg was the driver behind the epidemiological description and classification of RA [27]. Consequently, during the last decades of the last century, clinical research has focused especially on treatment strategies, leading to new and innovative insights in the importance of disease activity monitoring and long term outcomes [28]. Unsurprisingly, the development and first clinical studies in RA of TNFi had a strong basis in the Netherlands [29]. After the reimbursement of the first TNFi in the Netherlands, infliximab, by prof. dr. Els Borst the call for real-world evidence on the effectiveness and safety of biologicals like TNFi was loud and clear. She pursued the cause of Evidence-Based Medicine (EBM) and argued that incorporating Real-World Evidence (RWE) is instrumental due to its generalizability and more comprehensive nature. Additionally, RWE provides insight into healthcare utilization, patient adherence and long-term outcomes, all important factors surround the introduction of biologicals. On the initiative of prof. dr. Piet van Riel, research hospitals in the east of the Netherlands, have collected prospective data from consecutive RA patients for over 20 years. This evolved into the Dutch Rheumatoid Arthritis Monitoring Registry (DREAM). The DREAM registry has produced dozens of peer-reviewed publications on effectiveness, safety and efficiency of various treatment strategies, subgroups of patients and predictive characteristics [30].

Ultimately, this registry was not limited to RA, data was also collected from other rheumatic diseases. The DREAM registry prospectively captured not only medical data but also patient characteristics, preferences and outcomes as well as treatments, complications and costs. Contrary

to most other registries, DREAM focussed on patients as a whole from the moment of diagnosis, capturing full treatment strategies where others were focussing on specific interventions. This resulted in strong Real World Evidence for treating to the target of remission (T2T) methods in RA as well as the (cost-)effectiveness of different strategies for starting and stopping biological and targeted synthetic DMARDs. Obviously, the DREAM registry allowed the extension of observed Real World Data by modelling, contrasting modelling based on data form randomized controlled trials. One of the most recent studies in the history of DREAM is the PERFECTRA study, described hereafter in this thesis. Again, this encompassed unselected patients, in a pragmatic study providing real-world data (RWD) and RWE on a pertinent clinical and health-economical question with respect to which strategies help patients best at what cost.

Relevance

RA can be a prevalent, painful, disabling systemic auto-immune disease. RA can be classified as a condition that comes in high volume at high costs. The knowledge of the pathology and treatment options has increased enormously over the past decades. Numerous effective pharmacological treatments have been developed, and new strategies have been introduced over the past decades. Many treatments used in rheumatology, bDMARDs and JAKi, belong to the group of expensive drugs and the consequences of RA seriously burden not only the patient but also society as a whole. Approximately 270.000 people live with RA in the Netherlands. More than 10.000 patients are diagnosed each year [31]. In 2019, €292 million was spent on RA care. This accounts for 0.3% of all healthcare spending in the Netherlands. In comparison with more recent pharmaceutical innovations, RA medications like TNFi and JAKi do not account for the top 25 most expensive treatments currently available in the Netherlands when considering cost per patient, total expenditure on adalimumab (a TNFi) was ranked 9th highest total expenditure in the Netherlands by the Dutch Healthcare Authority in 2022 [32]. It should be noted that this might not be solely due to use in RA, as adalimumab is also indicated for multiple other rheumatic diseases, gastrointestinal indications, and eye infections. Rheumatology is one of the fields from which drug repurposing has been successfully conducted. Due to the chronic nature of RA, lifelong disease management is necessary, which leads to significant expenditure over time. Therefore, health economic analyses are important to support decision-making. The vast amount of options and possible treatment strategies and sequences for the treatment of RA provides many choices and routes. At the introduction of new interventions, it is not always immediately clear where this novel approach should fit in the grand scheme of strategies and options. Long-term outcomes, external validity, and real-world treatment patterns are instrumental in understanding medical, economic, and societal implications of new interventions. Health economic analyses can provide context and guidance. This will be elaborated further in upcoming sections.

Gout is an ancient metabolic disease, increasingly prevalent due to heightened global welfare. This high volume disease has only limited effective treatment options, which is why the societal impact of gout is to be considered seriously. In the Netherlands, due to the aging population, the amount of newly diagnosed gout patients is increasing. In 2017, there were 477,000 patients with gout, in 2030 this number is expected to rise to 580,000 [33]. Future innovative gout treatments are to be expected to come with considerable costs. In order to make decisions on these and the broader

context of gout treatment we need information to make informed decisions. The combination of high costs and high volume in gout treatment requires not only the attention of medicine but of health policy makers and health economists as well.

Pharmacological interventions

Registration and licensing of (pharmacological) innovation is based on pre-clinical and well-controlled pivotal trials, only including highly selected patients, excluding children and elderly as well as patients with comorbidities and/or using comedication. Promising pivotal trial data need to be confirmed in real life, where data can only be captured after market authorization. Obviously, RWD collection has a reduced level of control than pivotal trials and harnesses uncertainties with respect to diagnosis, prescription, daily use, and monitoring. Additionally, comorbidities and co-medication, and known or unknown risk factors, will result in variability introduced in terms of efficacy or observed side effects. Thus, waiting for real world experience, registry data, is on the one hand extremely time-consuming and not always without methodological concerns, but on the other hand provides information on risks and benefits of interventions after introduction in the market.

In the absence of RWD, the need to predict the future is increasingly pressing for health policy makers and administrators. Modelling, using available data in mathematical algorithms in order to predict the future has vastly proven its value outside healthcare. A theoretical framework firstly to understand the healthcare system and disease subject at hand, and secondly to help to program these algorithms is necessary. Economic models enable the simulation of situations and disease processes very accurately, even when including consumer (or patient) behaviour. The knowledge and experience with these models is increasing every day, with new abilities following suite. For instance, the capability to include disease activity, patient-specific characteristics, and effectiveness of treatment to very accurate levels. This further allows health economists to make correct estimates of costs, effectiveness and the created value of not only singular treatments but also treatment strategies and the larger context in which they are given.

Health Systems Value Assessment

Health Economy (HE) is servicing health policy makers. On the one hand, HE can help to understand and organize healthcare systems and its financial backbone. Ideally, this should be data driven. Information on costs, effects, and societal impacts are the basis on which decisions can be made. However, taken into account the diversity, complexity, and magnitude of healthcare systems with continuing promising innovations in rheumatology as well as in other medical fields, reliable data sets are not always available or hard to be analysed reliably. Additionally, there are many more factors apart from pharmaceutical innovations that drive healthcare expenditure. Aging populations, increased demand for care, advances in medical technology, and workforce changes all factor into expenditure and policy decision. In general, politicians tend to focus on budgets, budget impact, and budget restrictions, rather than on increased created value. Even years after his introduction of value-based healthcare, we must conclude, probably not surprisingly, that there is a long way to go before healthcare is completely based on created value.

Healthcare is one of the largest and fastest-growing industries. High-income countries spend a large portion of their Gross Domestic Product (GDP) on healthcare [34] - and these numbers are growing. Table 1 displays current health expenditure (CHE) as a percentage of GDP in the Netherlands and some surrounding countries as presented by the World Health Organization (WHO) Global Expenditure Database [35]. With the exception of Luxembourg, all countries spent more than 10% of their GDP on healthcare. Naturally, exact definitions of healthcare spending can vary per country.

Table 1 Healthcare expenditure as a percentage of a countries' Gross Domestic Product by World Health Organization Global Health Expenditure Database.

Countries	2015	2020	2021
Belgium	10.80	11.20	11.04
Denmark	10.34	10.56	10.82
Finland	9.65	9.63	10.25
France	11.45	12.13	12.31
Germany	11.19	12.69	12.93
Luxembourg	5.08	5.74	5.67
The Netherlands	10.32	11.21	11.29
New Zealand	9.28	10.07	10.05
Norway	10.07	11.42	10.08
Sweden	10.80	11.33	11.25
Switzerland	10.78	11.73	11.80
United Kingdom of Great Britain and Northern Ireland	9.80	12.16	12.36
United States of America	16.49	18.76	17.36

The high priority of health and healthcare by populations in high-income countries, in combination with the ongoing innovation and technological progress have been credited as main drivers for the rising spending on healthcare [36]. The healthcare industry, from an economist's perspective is unique, with phenomena like the productivity paradox also affecting (perceived) efficiency and thus the cost of healthcare [37]. Not surprisingly, in the past decades, increasing concerns surface on the budget impact of these innovative treatments. Moreover, health policy makers are confronted with decisions to be made on promised innovations, without reliable information on budget impact and (long-term) effects on health status. This becomes increasingly difficult when these decisions need to be made for small highly selected populations, where classic formal randomized controlled trials (RCTs) are nearly impossible due to small patient numbers worldwide. These decisions do not solely regard implementation of innovations of treatments but also relate to disinvestments and active reconsideration of already implemented innovations. In every society, history has proven that stopping an implemented innovation results in disappointment and protest. Even when the balance between costs and gains for an innovation are unfavourable. Especially in healthcare this mechanism occurs. Disinvestments should be very clearly supported by evidence and explanations to those affected by the decisions made. Making choices in the healthcare system is inevitable so it is of great importance that decision makers have a solid system in place in order to do so. If we solely use medical considerations for these decisions, there will be

and unbridled growth of costs. Societal and economic impacts have to be considered as well. Justice, transparency, and ethics should also be considered when making these decisions.

The widely accepted framework provided by prof. Michael Porter and prof. Elizabeth Olmsted Teisberg provides a way of assessment of interventions in which not cost, but added value should be the driver for decision making [38]. They argue that not the service or acts, but the outcomes of the patient's treatment should be the determinant of value. Thus, improvement of health status divided by cost is what defines value. A value that, ideally, can be compared across providers and conditions. In order to assess such improvement of health status, it is important to be able to appropriately assess what is value and how and which outcomes should be measured. On the other hand, while cost may seem an easier dimension to measure, accurate reporting and assessment of all costs associated with health technological interventions. In contrast to the present situation, it is imperative that society, health policy makers, and patients focus on creating value rather than on structures, processes and academic prestige. Transparency of outcomes helps patients in their decision-making and helps payers to see whether their funds are allocated and spent effectively. Also, health policy makers can use outcome transparency for strategic decisions. Therefore, value-based healthcare is data-driven healthcare. Data on outcomes that matter to patients determine the created value. This is where real-world data becomes real-world evidence.

Obviously, in the case that complete data sets are available it seems easy to choose and decide. However, decisions often must be made in suboptimal scenarios where not all desired data is available or of the appropriate quality. One option is to perform sensitivity analyses, in which the uncertainty around certain parameters can be assessed and taken into account when analysing. Another option is the use of available data from different sources and combining those data in a methodologically correct way creates opportunity to make these necessary decisions in the best-informed manner possible. These difficult decisions should be made with the most available resources, rather than based on opinions or convictions.

Most innovative treatments being brought to market in the Netherlands will fall under the category defined as *expensive treatments*. In the Netherlands, expensive treatments are defined as treatments that cost more than €1.000 per patient, per year [32]. In the field of rheumatology, this means that most treatments fall into this category. Expensive treatments are responsible for a large portion of total healthcare spending and that percentage increases yearly. This is a group of medications and treatments in which hospital pharmacists, healthcare insurers, medical professionals, and pharmacies allocated vast resources to ensure the best quality of care along with accessibility of healthcare for all members of society. New and innovative treatments becoming available for (groups of) patients is incredibly important and can change lives of patients and their families. As science is progressing and more treatments are being introduced, it is vital to assess efficacy, effectiveness, cost-effectiveness. It is vital to ensure the comparability of these studies, in terms of methodology, (collection of) cost information, quality of life assessment, and other measures that matter to patients as well as those that are clinically relevant.

Health Technology Assessment

The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) has provided a comprehensive approach to Health Technology Assessment (HTA) evaluating multiple dimensions in order to assess added value of interventions [39,40].

The ISPOR HTA Dimensions framework concerns:

1: Current Use, 2: Technical Aspects, 3: Safety Aspects, 4: Clinical Effectiveness, 5: Cost and Economic Aspects, 6: Ethical Analysis, 7: Organizational Aspects, 8: Patient and Social Aspects, 9: Legal Considerations

Current use (1) relates to the indication or disease that is evaluated and how it is currently managed and treatment. This also relates to incidence or prevalence, providing a first idea of impact. Technical aspects (2) explore the treatment pathway, mode of action, or other distinguishing features of the studied intervention. Safety aspects (3) evaluate the safety of the intervention, relating to adverse events and other potential risks. Clinical effectiveness (4) relates to the assessment of effectiveness of the health technology or pharmacological intervention towards its intended endpoint.

Points 1-4 should be taken into consideration for rapid relative effectiveness assessments. The full framework, points 1-9, make up the Full HTA framework dimensions.

Costs and economic aspects (5) examine budget impact, cost-effectiveness, cost savings, or broader impact on the healthcare system. Ethical analysis (6) should take into account equitable access, resource allocation, and fair distribution. Organizational aspects (7) explore the impact of the studied intervention on hospitals and other care providers, technical operational aspects of reimbursement, and other stakeholders that might be impacted by the studied intervention. Patient and social aspects (8) should focus on the impact of the studied intervention on the patient in a broad sense. Legal considerations (9) allow analysis of the legal implications of the studied intervention, possible bottlenecks, or regulations that need to be in place.

This dissertation applies (parts of) the full HTA framework to assess treatment strategies in rheumatology. This dissertation focusses on current use, clinical effectiveness, cost & economic aspects, organisational aspects, and patient & social aspects through cost-effectiveness models, analysis of a pragmatic real-world setting trial, and an assessment of responsiveness of disease-specific and generic measures of patient's health status. The ISPOR Core HTA model provides a structured approach to evaluation of treatment strategies in rheumatology.

Health Economic Methods

Through a series of peer-reviewed studies, this dissertation utilizes health economic methods to provide insight into cost-effectiveness of various treatment strategies. Markov models, combinations of multiple data sources and novel health state definitions are among the main concepts utilized. Markov modelling is a technique that was originally employed towards the modelling of stochastic processes, brought to light by mathematician Andrey Markov. It has since proven useful in various applications, due to its versatility. In healthcare and health economics,

modelling has proven necessary for economic evaluation. As Drummond et al [41] discussed in 1997, although trial-based health economic evaluations are the preferred way, modelling is necessary when not all and any data is captured and available. This still holds true today. Markov modelling is a widely used tool for economic evaluations in healthcare and can, for example, be well suited when simulating chronic disease processes, like RA or gout [42]. In Markov modelling, a variable that changes through time is followed. In health economic modelling, these states are often referred to as health states. Markov models allow the incorporation of costs, effects, and other characteristics linked to those health states, making this method very useful for economic evaluation.

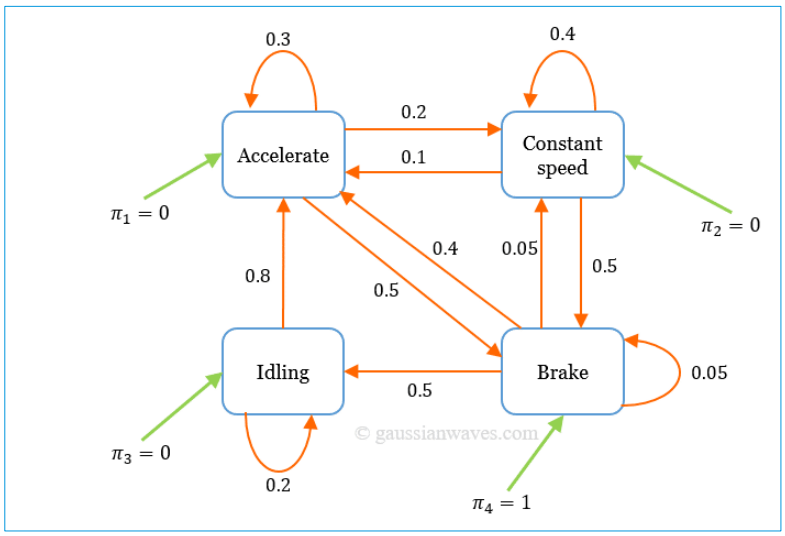


Figure 7 Example of Markov chain transition model

Naturally, there are some limitations. The Markovian assumption, on which the model is built, assumes that the probability of moving from one (health) state to another does not depend on any states that the model was in previously. In other words, it assumes that the current probability of movement to another health state for a given simulated patient does not depend on any health state this simulated patient has held in the past. Naturally, when simulating disease processes, it is not always this simple. As described by Briggs and Sculpher [42], there are ways to deal with this and incorporate more flexibility into a good model, so that models simulate disease processes very closely. Obviously, when using modelling to simulate disease processes, one has to account properly for uncertainty. Deterministic analysis executes the model at parameter means, giving a single output for decision-making. However, uncertainty surrounding model input should be considered. Probabilistic sensitivity analysis is a tool through which parameter uncertainty in (Markov) models can be represented and addressed [43,44]. Markov models are non-linear, the results of probabilistic analyses will not necessarily match the deterministic analysis. When utilising a Markov model for health economic evaluation, uncertainty will be present on various input parameters such as transition probabilities, costs, and utility. This is due to the variability in the sources from which data is retrieved. Probabilistic sensitivity analysis accounts for the uncertainty by fitting distributions individually around all inputs. Then, many simulation rounds are conducted which leads to a distribution of outcomes [45]. Presenting a distribution of

outcomes in a way that is readable and understandable not only for health economists but also for decision makers with varying backgrounds is of pivotal importance. One tool to do so is the Cost-Effectiveness Acceptability Curve (CEAC) [46–48]. Originally brought forward as an additional option to the incremental cost-effectiveness ratio (ICER) [49], it shows the probability that the studied intervention is cost-effective at certain willingness-to-pay (WTP) thresholds.

At the basis of health economic evaluation, it is necessary to collect the right information. This includes information on a patient's health status and clinical outcomes. This can be captured by, for example, the EuroQol 5 domain (EQ-5D) [50] or the Health Assessment Questionnaire Disability Index (HAQ-DI) [51]. When introducing or evaluating health-related Quality of Life (HRQOL) measures, assessing responsiveness is one of the major factors [52,53]. Responsiveness relates to the accuracy with which a certain measure captures an actual change. In this dissertation, responsiveness in case of a deterioration in health status was assessed. This is an important distinction. Sensitivity to change of HRQOL instruments in case of improvements does not necessarily translate to sensitivity to change in the case of deterioration. Nonetheless, this is of great importance due to the need for response in cases where a patient experiences a decline in health state. Additionally, data that is suitable for responsiveness assessment where there is a decline in health status is not widely available, as, naturally, most trials and interventions are aimed at improving patient’s health.

In the field of rheumatology, over the past 25 years, many innovative pharmacological interventions are developed and introduced to the market. All these interventions came at considerable prices, on the one hand driving the costs of pharmacological interventions while on the other hand reducing the burden of the disease of patients. Where many rheumatic diseases, for instance RA, almost certainly lead to disability, impairment of workability and high dependency of care, at present patients suffering RA have the option of cure and frequently even drug free remission. For HE this is a very interesting field to explore. Not only how different systems organise and pay for such innovation, but also how they use data to analyse whether or not the created value reflects well-allocated and spent resources. In the Netherlands, rheumatology succeeded to prospectively capture data in the full picture of medical, patient reported and costs. Especially the DREAM registry was considered a best practise for RA and other conditions. This thesis uses the opportunity of the DREAM registry to perform HE analyses with this data set and use the Real World Data for various modelling exercises. The results of this research have been published in peer reviewed or submitted papers that form the subsequent chapters.

Outline of the Thesis

This research explores different aspects of health economics in the field of rheumatology. All research topics make strong use of real-world data towards information generation on cost-effectiveness, clinical effectiveness, and measurement considerations. A clear focus lies on the decision-making in treatment strategies with regards to the guidelines for treatment of rheumatic diseases. Research topics include the ISPOR Full HTA Core Model subtopics current use, clinical effectiveness, and cost & economic aspects.

Chapter 2 assesses the cost-effectiveness of two full treatment strategies used in daily clinical practice through a Markov model cost-effectiveness analysis. The question lies in step-up therapy versus a more intensive initial combination therapy. Cost-effectiveness of these two real-world implemented strategies is modelled. **Chapter 3** combines real-world data with data from RCTs to analyse the cost-effectiveness of an implemented treatment strategy to one where a key treatment in that strategy is replaced by an innovative treatment.

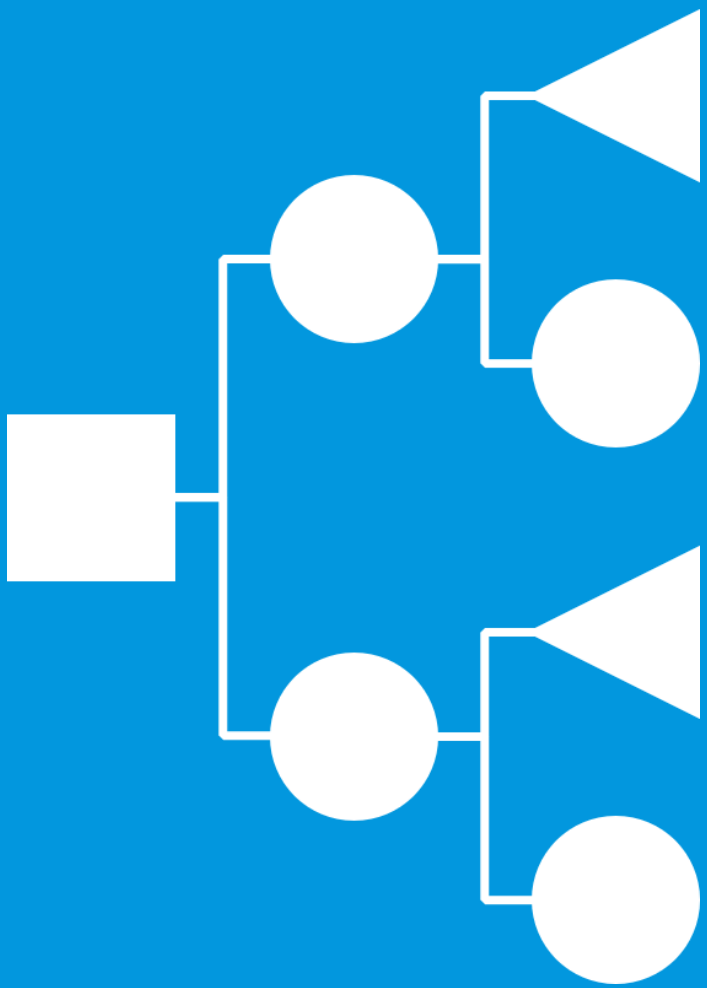
In chapter 4 we answer the question of which strategy of gout management is cost-effective through a model in which a novel definition of gout health states was utilized. Both the underlying disease mechanism and gout flares require medical treatment, which leads to a vast amount of treatment strategy options.

Chapter 5 explores the responsiveness of widely-used measures of HRQOL, in case of deterioration in health state through analysis of a unique dataset in which patients discontinue treatment. HRQOL measures responsiveness assessment, in improvement and deterioration, is key towards proper inputs in health economic analyses.

Chapter 6 answers the question of clinical effectiveness of two treatment strategies in RA treatment in a real-world setting. The PERFECTRA study is an open-label pragmatic real world study in adult patients with active RA who did not respond (well) to the first line of treatment, conventional DMARDs assessing real-world effectiveness as a supplement to pivotal RCTs.

Chapter 7 contains a general discussion covering the aforementioned themes.

Section I



Section I - Different Health Economic Models in Rheumatology Treatment

Chapter 2

Chapter 2 – New Markov Model for Rheumatoid Arthritis Disease Simulation

Published as: Cost-effectiveness of different treat-to-target strategies in rheumatoid arthritis: results from the DREAM registry

Van de Laar CJ, Oude Voshaar MA, Vonkeman HE. Cost-effectiveness of different treat-to-target strategies in rheumatoid arthritis: results from the DREAM registry. BMC rheumatology. 2019 Dec;3:1-9.

Abstract

Background

Adjusting medication of patients with rheumatoid arthritis (RA) until predefined disease activity targets are met, i.e. Treat-to-Target (T2T), is the currently recommended treatment approach. However, not much is known about long-term cost-effectiveness of different T2T strategies.

We model the 5-year costs and effects of a step-up approach (MTX mono -> MTX + csDMARD combination -> Adalimumab -> second anti-TNF) and an initial combination therapy approach (MTX + csDMARD -> MTX + csDMARD higher dose -> anti-TNFs) from the healthcare and societal perspectives, by adapting a previously validated Markov model.

Methods

We constructed a Markov model in which 3-monthly transitions between DAS28-defined health states of remission (≤2.6), low (2.6<DAS28≤3.2), moderate (3.2<DAS28≤5.1), and high disease activity (DAS28>5.1) were simulated. Modelled patients proceeded to subsequent treatments in case of non-remission at each (3-month) cycle start. In case of remission for two consecutive cycles medication was tapered, until medication-free remission was achieved. Transition probabilities for individual treatment steps were estimated using data of Dutch Rheumatology Monitoring registry Remission Induction Cohort I (step-up) and II (initial combination). Expected costs, utility, and ICER after 5 years were compared between the two strategies. To account for parameter uncertainty, probabilistic sensitivity analysis was employed through Gamma, Normal, and Dirichlet distributions. All utilities, costs, and transition probabilities were replaced by fitted distributions.

Results

Over a 5-year timespan, initial combination therapy was less costly and more effective than step-up therapy. Initial combination therapy accrued €16226.3 and 3.552 QALY vs €20183.3 and 3.517 QALYs for step-up therapy. This resulted in a negative ICER, indicating that initial combination therapy was both less costly and more effective in terms of utility gained. This can be explained by higher (±5%) remission percentages in initial combination strategy at all time points. More patients in remission generates less healthcare and productivity loss costs and higher utility. Additionally, higher remission percentages caused less bDMARD use in the initial combination strategy, lowering overall costs.

Conclusion

Initial combination therapy was found favourable over step-up therapy in the treatment of Rheumatoid Arthritis, when considering cost-effectiveness. Initial combination therapy resulted in more utility at a lower cost over 5 years.

Background

Rheumatoid arthritis (RA) is a systemic autoimmune disease with alternating periods of lower and higher disease activity. RA may have a chronic, progressive course, leading to functional impairment and reduced quality of life [54]. The main objective of treatment is to achieve suppression of inflammation as soon as possible, to minimize symptoms in the short-term and to retard progression of structural damage in the long term. The adoption of modern treatment strategies, together with new and expensive, biological or targeted synthetic, disease-modifying anti-rheumatic drugs (DMARDs) have considerably improved patient outcomes. However, the cost of these new drugs combined with the lifelong scope of RA treatment has resulted in a considerable cost burden on payers of healthcare costs [55].

The approach currently recommended for RA treatment involves titrating medication dosages until pre-specified disease activity targets (both remission and low disease activity (LDA) or LDA) have been met and maintaining these targets over time. Such so-called treat to target strategies (T2T) have proven to be more effective and to generate more utility than usual care [56,57]. A previous study by Vermeer et al. found that the focus on rapid suppression of inflammation results in high initial costs, but has been shown to be well within willingness-to-pay thresholds in the long run [56]. The European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) have developed comprehensive recommendations on the setup and implementation of T2T in clinical practice [58,59]. However, even when following these guidelines, different treatment strategies can be adopted, for example T2T protocols employing step-up therapy, initial combination therapy or initial biological DMARDs therapy. These differences may lead to important variation in clinical outcomes, costs, and utility.

Not much is known about real world cost-effectiveness of such alternatives however, because multiple previous health-economic evaluations of RA treatment strategies focus on the optimal place of one particular drug in some sequence of treatment options [60]. Moreover, models are typically fed with data from clinical trials with selected patient populations. This limits the generalizability of the results and could potentially wrongly estimate real-world treatment effectiveness for various reasons, such as selection criteria of trials favouring patients likely to respond, and wash-out period before treatment initiation [61].

The aim of this paper is to compare the long-term cost-effectiveness of step-up therapy and initial combination therapy from a societal perspective, by expanding on a previously validated Markov modelling approach and populating the model using data extracted from two real world cohorts of unselected RA patients that have been treated using the respectively modelled strategies. This will allows to improve the conceptual model framework for RA which can be used in the future for more comparisons of treatments and strategies.

Materials and Methods

Health economic modelling

This study used a Markov model to assess long term cost-effectiveness of two T2T strategies in treatment of rheumatoid arthritis. Step-up therapy was compared to initial combination therapy

in a model based health economic evaluation in which economic consequences of two treatment strategies are evaluated within a mathematical framework. In accordance with the ISPOR principles [62] for good practice for decision analytic modelling in healthcare evaluation, all model input for our study was derived directly from the various DREAM cohorts, as described in the next section.

Data

All data used in this study was derived from two real world observational studies in which patients were treated according to T2T protocols, both aiming at achieving 28 joint-count disease activity score (DAS28) remission (i.e. DAS28 score ≤2.6) in order to extrapolate over a period of 5 years. Baseline characteristics of the patients are summarized in supplemental table 1.

Outcomes and costs were registered in the same way in both cohorts and data collection, including DAS28-assessment, was carried out by trained rheumatology nurses. Patients were included upon diagnosis with early-onset moderate to severe RA (DAS28>3.2) and were DMARD-naïve. For both cohorts, data was collected in the same hospitals.

Patients in Remission Induction Cohort I (RIC I) (step-up therapy) were initially treated with methotrexate (MTX) monotherapy, followed by addition of sulfasalazine. In case of persistent moderate disease activity (moderate or high; DAS28 >3.2) sulfasalazine could be replaced by TNFi. Due to reimbursement policies in the Netherlands, patients with DAS28 >3.2 were allowed to start TNFi. Patients were evaluated at baseline, 8, 12, 20, 24, 36, and 52 weeks, and every 3 months thereafter [63]. Consecutive patients entered this cohort between 2006 and 2012 and were followed regularly thereafter.

Patients in Remission Induction Cohort II (RIC II) (initial combination) were initially treated with combination csDMARD therapy, followed by high-dose combination therapy. In case of persistent moderate to high disease activity (>3.2), a TNFi could be started, replacing one of the csDMARDs. Patients were assessed at baseline, 2, 4, 6 months and every 3 months thereafter [64]. Consecutive patients entered this cohort from 2012 onward and were followed regularly thereafter.

Patients in both cohorts have given written informed consent before inclusion. The attending physician and the patients were advised to follow the per-protocol predefined assessments and treatment decisions. Treatment changes could be made at any time point at the discretion of the rheumatologist. In general, conformity to the protocol was good [65].

Model input

Transition matrices for each treatment step were derived from data obtained from the subpopulation of patients treated with the relevant medication and dosages. Chi-squared statistics tests were used to verify the stability over time of the obtained transition probabilities [66]. All DAS-28 measures obtained during the period of time patients were treated using medication and dosages relevant to a treatment step were used to estimate 3 monthly transition probabilities from the sample proportions [67,68]. Since clinic visits were not always scheduled in exact 3 month increments, a range of 1.5 months was used for DAS28 measurement moments. The distribution

of patients of the four DAS28 states will be compared to observed daily clinical practice outcomes in RIC I.

The EuroQol five dimensional (EQ-5D) [69] questionnaire was used to value the quality of life in all four respective health states. The average EQ-5D utility score of observed patients in each of the four DAS28 health states is used in the model.

The EQ-5D was recorded during all clinical visits. Utility scores of patients in each of the four DAS28 states were averaged. The questionnaire assesses a patient's well-being on five dimensions: mobility, self-care, daily activities, pain/discomfort and anxiety/depression. Each dimension is valued using three levels: no problems, some problems, or extreme problems. The EQ-5D disserts $3^5 = 243$ different states of health. Each state is valued to create a utility score between 0 and 1 using the Dutch EQ-5D tariff [70]. A score of 0 represents a state that is as desirable as death. A score of 1 is considered perfect health.

To accurately reflect the costs that are connected to specific health states, both healthcare consumption and cost of absence from paid labour are included. A healthcare consumption questionnaire with questions about the type and amount of different care 'units' patients have received since their last appointment with the rheumatologist was used. This includes appointments with any type of specialist, general physician, or use of other types of care or medication. These care units were multiplied using the 2016 updated price index numbers to calculate the average amount of care consumption per health state [71]. Average sick days are measured as the average number of days per 3 months that patients have reported sick to their employers. It was assumed that the cost of one employee not being able to work for 1 day is €230 based on a report by The Netherlands Organisation for applied scientific research in 2014 [72]. To account for the fact that not all patients have paid jobs, the proportion of paid jobs patients (split by age and sex) is multiplied with average sick days. Leading to

Equation 1

$$SD_{ij} * \text{workforce participation} * 230 + ZoCo_{ij}$$

(i and j respectively health state and cohort)

for total health state-specific cost in both cohorts. Where SD refers to number of sick days per state and ZoCo refers to the Dutch Healthcare Consumption Questionnaire (HCQ).

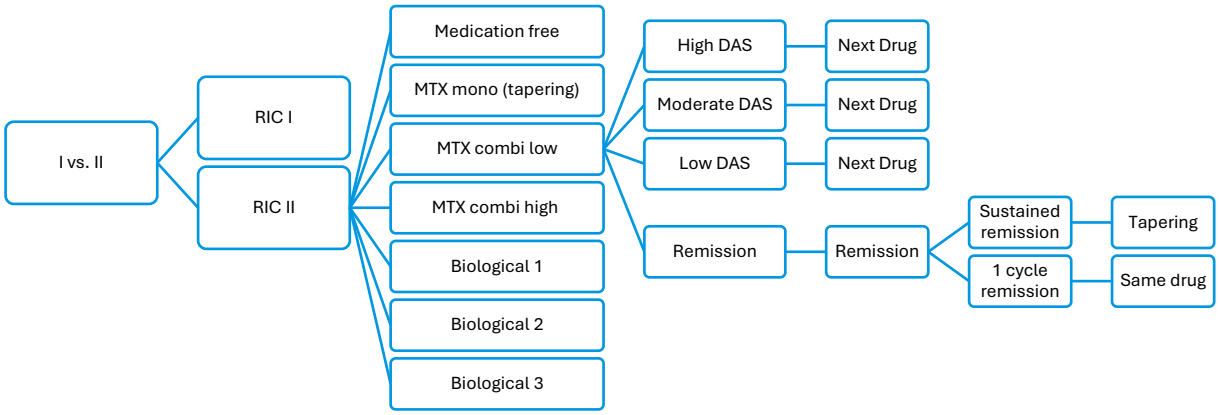
Markov model

Markov models can synthesize different types of costs and outcomes (utility, effectiveness) over a specified time [73–75]. For this study, using the data from DREAM RIC I and II, 5-year outcomes of two T2T-strategies were simulated using such a Markov model. Models consist of a finite number of 'Markov states' through which modelled patients move, with the probability to move from one state to another depending on transition probabilities. Different costs and outcomes accrue and lead to the eventual quality adjusted life years and healthcare costs, depending on the Markov states and treatments modelled patients move through.

In the Markov model that was constructed for this study, patients are always in one of four mutually exclusive disease activity related health states. The health states are defined by the commonly used disease activity score in 28 joints (DAS28): remission (DAS28≤2.6), low disease activity (2.6<DAS28≤3.2), moderate disease activity (3.2<DAS28≤5.1), and high disease activity (DAS28>5.1) [76]. A time horizon of 5 years was used. This time horizon is divided into 20 cycles of 3 months and modelled patients may shift from one health state to another at the start of each cycle, with the transition probabilities depending on their health state at the beginning of the cycle and the medication they are using at the start of that cycle [73]. Due to the fact that in real life, transitions are not automatically expected to occur at the beginning of a 3-monthly cycles, within-cycle correction was applied. Patients could move to a different health state at any point in that cycle. This method corrects cycle rewards and cost overestimation by considering the percentage of patients in each health state at the beginning and end of the cycle.

All patients initially enter the model on the first medication of their treatment protocol (resp. MTX monotherapy (RIC I) or MTX combination therapy (RIC II)). See figure 1. After the first cycle, patients will either be in remission (DAS28≤2.6) and stay on the same drug for another cycle, or not in remission (DAS28>2.6) and progress to the next drug as prescribed by the protocol. When patients sustain remission for two cycles (6 months), their medication will be tapered, as specified by the protocols. They will move to the preceding drug, or a medication-free state, in case no more preceding drugs are available and if their remission sustains for an adequate amount of time. For example, a patient that sustains remission on MTX monotherapy (in RIC I) for 2 cycles, will ‘jump’ to low-dose MTX. If the patient sustains remission for another six months, he/she will ‘jump’ to the medication-free state. Modelled patients will move one medication step up in case of a flare (DAS28>2.6), until it the flare under control.

Figure 1: Model structure



This figure represents the Markov Structure of the initial combination strategy. MTX combi low: low dose csDMARD combination therapy; MTX combi high: High dose csDMARD combination therapy; Biological 1,2,3: bDMARD therapy. High DAS: High DAS28, Moderate DAS: Moderate DAS28; Low DAS: Low DAS28; Next Drug: patients move downstream; Remission: in case of sustained remission medication will be tapered, in case of only 1 cycle in remission, patient continues the same treatment.

Outcomes

All outputs in the model are globally discounted annually, using 4% for cost and 1.5% for effectiveness, as recommended in the Dutch Guideline for Economic Evaluations in Healthcare [71]. The primary outcome measure was the incremental cost effectiveness ratio (ICER). i.e., the incremental cost for one additional quality adjusted life year (QALY). In the Netherlands, the generally accepted threshold below which treatments are considered cost-effective lies between €60,000 and €100,000. In this paper, both extremities will be considered.

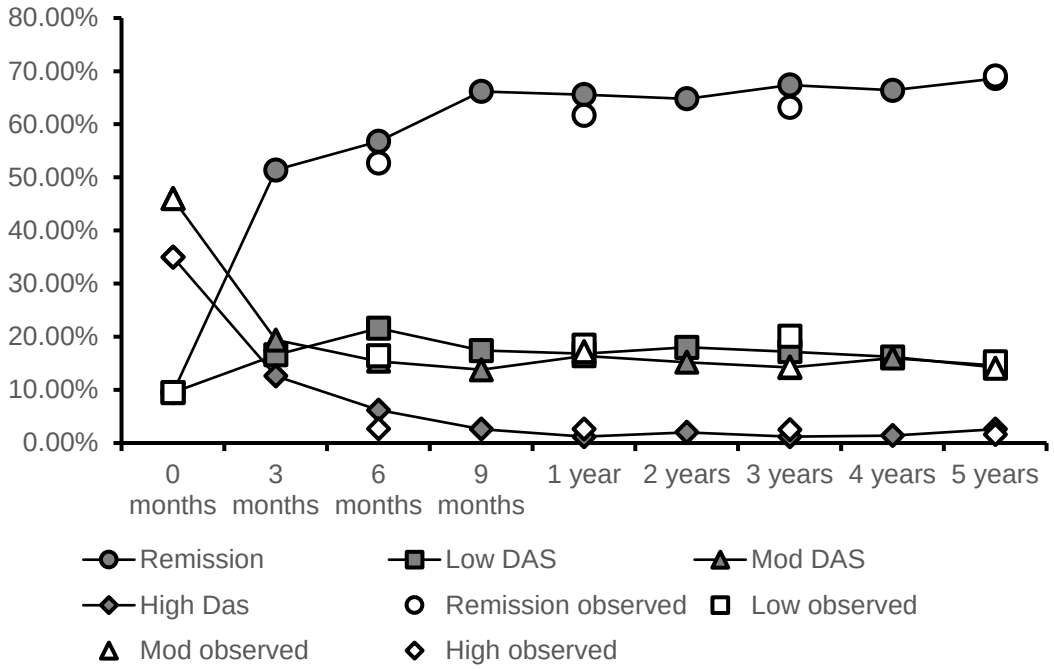
Equation 2

$$ICER = \frac{\Delta Cost}{\Delta Effectiveness} = \frac{C_{II} - C_I}{E_{II} - E_I}$$

where the subscripts I and II refer to the compared interventions.

Monte Carlo Simulation

5000 patients were simulated individually using Monte Carlo simulation, which allows to keep track of the disease course of each modelled patient as they moved through different cycles. The model was constructed and analysed using TreeAge Pro software (Williamstown MA, USA). The software assigns and records all Markov states, transitions, and costs and outcomes to each individual patient.



Probabilistic Sensitivity Analysis

To account for uncertainty in the parameter estimates, distributions for all cost and utility estimates were fitted. Different distributions were evaluated in order to find the most appropriate ones for these parameters. Different suggestions and guidelines were evaluated [77–79]. These included normal, gamma, logistic, beta, and Poisson distributions. Using Anderson Darling/SK and Chi^2 tests, the most appropriate distributions were selected. The transition probabilities also face a level of uncertainty. All transition probability matrices have been re-specified using Dirichlet distributions. The matrices that included less data will incorporate more uncertainty than those with higher numbers of observations. In the probabilistic sensitivity analysis, 2000 runs of the model were performed where the input parameters were re-sampled from these distributions for each iteration.

Results

Comparison of model predicted- and clinically observed disease activity outcomes over 5 years

Figure 2: model validation

Figure 2. Distribution of simulation cohort over DAS-28 states, compared observed distributions of RA patients in RIC I. Observed data is available at baseline (corresponds to baseline distribution for modelled patients), 6 months, 1 year, 3 years and at 5 years.

Figure 2 presents the distribution of the modelled patients over the four DAS28 states for each of the twenty 3-monthly cycles, compared with the distribution of patient over the disease activity states as actually observed in daily clinical practice in RIC I [63]. The percentage of modelled patients in remission increases from ~10% at baseline to about 65% at the 1 year visit, which closely corresponds to the remission percentages seen in the cohort in the actual patients. Similar results can be seen for the other disease activity states, which supports the validity of the projected disease activity outcomes that are generated by the model. Comparison yields similar results when comparing the result of the initial combination strategy with result from RIC II.

Base analysis

The ICER of initial combination therapy versus step-up therapy with 5000 simulated patients is -139.000.Initial combination therapy is dominant over step-up therapy. . In this case, initial combination therapy is less costly (€20856.56 vs €25377.01) and more effective (3.54 vs 3.50 QALY) over 5 years. This indicates that initial combination therapy is cost-effective and dominates step-up therapy.

Table 1 Base Analysis results

	Step-up	Initial Combination
Mean costs (€)	25377.01	20856.56
Mean utility (QALY)	3.501	3.545
ICER (€)	-	-139,000

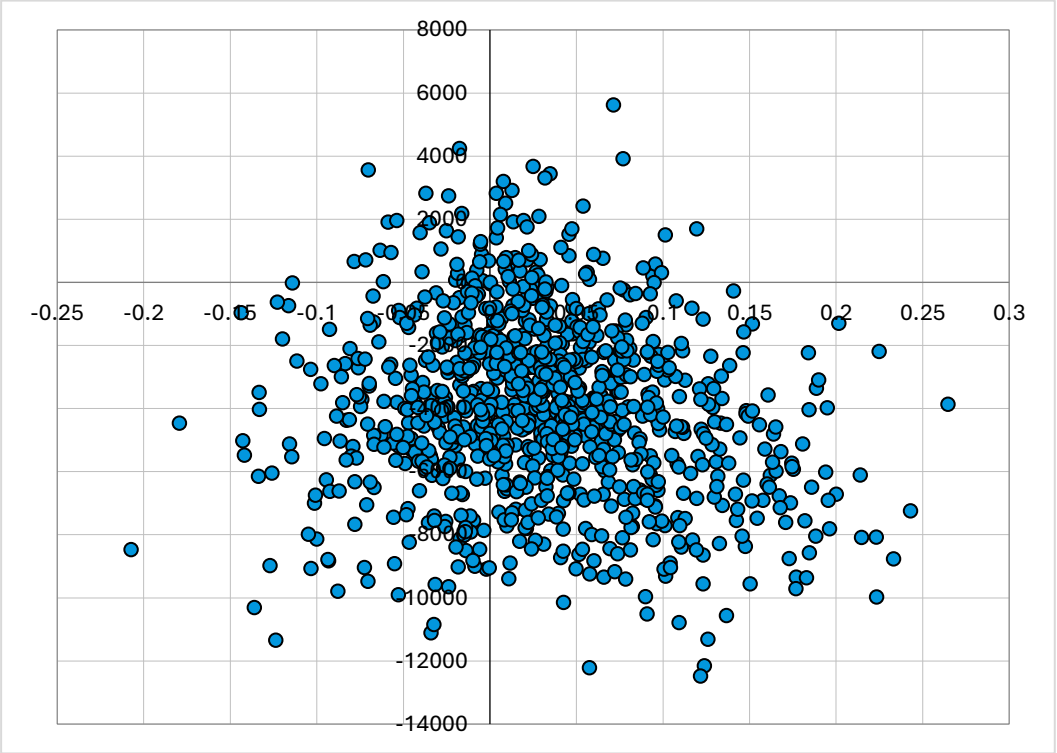
Table 2 depicts the results of the PSA. 2000 samples of 200 patients were run. The results show that initial combination therapy is cost effective, also when the uncertainty in the model inputs were considered in probabilistic sensitivity analysis. That strategy yields more QALYs but also saves cost in absolute terms. The difference in cost between the two strategies is almost €4000 over 5 years. The difference in accrued QALYs is smaller, 0.0325. Uncertainty around estimated costs and effects was small. The ICER is negative because initial combination therapy is a dominant strategy. These results were shown to be robust in the PSA. Figure 3 shows the ICER scatterplot for the probabilistic sensitivity results. This ICER plane shows the incremental cost-effectiveness of initial combination therapy strategy as compared to step-up therapy. It shows that a large proportion of trials is in the southeast quadrant, with positive incremental utility, and negative incremental costs.

Probabilistic Sensitivity Analysis

Table 2 Probabilistic Sensitivity Analysis results

	Step-up	Initial Combination
Mean costs (€)	€20163.81	€16267.15
2.5%-97.5% interval (€)	16588.46-23780.33	11534.72-21366.87
Mean utility (QALY)	3.515	3.548
2.5%-97.5% interval	2.467-4.598	2.44-4.71
ICER (€)	-	-119,897

Figure 3 Cost-effectiveness plane



Cost-effectiveness plane of the comparison of initial combination therapy with step-up therapy. Results of probabilistic sensitivity analysis. 500 patients displayed. X-axis: incremental effects (in Quality-Adjusted Life Years). Y-axis: Incremental costs (in €'s). 79% of trials are cost-effective, 63% are cost-saving.

Discussion

Treating to the target of remission of recent onset RA with combination therapy versus step-up therapy with disease modifying anti rheumatic drugs is more effective (in terms of EQ-5D utility) and less costly. According to the results in the current study, early-onset RA patients being treated

with initial combination therapy accrue higher utility and lower costs over five years as compared to patients being treated with step-up therapy. . A negative ICER is ambiguous, as it can indicate that the intervention is more costly while less effective (meaning it is dominated) or that it is less costly and more effective, hence, a dominant strategy. The negative ICER in this study shows that the second strategy, initial combination therapy, is both less costly and more effective, thus making it an absolute preferable option and dominant strategy over step-up therapy. The difference in accrued QALYs between the two modelled strategies is relatively small at 0.0325 over 5 years. This could be due to the fact that the strategies are both treating to the target of remission, thus have the same therapeutic goals.

Overall, there are more patients in remission in initial combination therapy strategy than in the step-up therapy strategy at each time point. This reduces health costs, increases utility, and medication can be tapered lowering medication costs. From a health-economic perspective, physicians should thus prefer initial combination over step-up therapy to treat to the target of remission strategy.

In this present study data obtained from two cohort studies in patients with early RA to project 5-year outcomes of two different treat-to-target strategies were used. Clinical outcomes obtained in these cohorts were previously compared [80], which showed more remissions at 6 months with initial combination therapy than step-up therapy. The results of this current modelling study suggest that this trend continues with higher remission percentages across all time points. The present study is among the first modelling studies to specifically evaluate outcomes of early RA patients who are enrolled in a tight-control strategy immediately upon diagnosis. In previous modelling studies in RA, it is usually assumed that the status of patients progressively deteriorates, which is typically modelled as decreasing Health Assessment Questionnaire Disability index (HAQ-DI)-scores over time. However, a pattern of deteriorating status is not consistent with studies describing 5-10 year outcomes of early RA cohort studies in which a tight control protocol was used, and was therefore not considered appropriate for the current study. In reports describing long term outcomes of such cohorts, HAQ-DI scores typically follow the same pattern as the DAS28 scores, with initial high disability followed by a prolonged (5-10 years) period of stable low disability according to HAQ-DI [81–87].

Schipper et al. first adapted the Welsing et al. Markov model [74] that was also used in his study. . In the paper by Schipper et al. [73], the model reaches an equilibrium at 2.5 years, with no further transitions occurring after that time point. This presumably happened because the majority of the patients were absorbed in the ‘sustained remission’ state. In the present study, the Schipper et al.’s [73] version of the model was adapted in several ways to better reflect the T2T based treatment strategies, as well as the disease course of RA, that is characterized by recurrent flares. Simulated patients were now able to reach sustained remission and have their medication tapered, but in case of a flare they were also able to return to their last effective medication. Remission was no longer an absorbing state and individual simulated patients’ disease course fluctuates over time, as in real-life. In line with studies on stopping and tapering of TNFi, in this model, biological-free remission has become an option for RA patients. Despite these adjustments, Schipper et al. have found comparable results over the 5-year modelling period. Accrued QALY’s over 5 years are

slightly lower, which could be explained by the different discount rate for utilities that the authors used. Schoels et al., [56] who performed a literature review of economic aspects of treatment sequences in RA have focussed on step-up therapy, similar to the one employed in this study, as the most cost-effective option versus employing TNF-i's in an earlier stage of the disease process. This paper confirms this notion and extends it to longer-term cost-effectiveness.

A major strength of this study is the use of real world daily clinical practice data from recent onset RA patients. The usage of daily clinical practice data assures that patients were not selected and the study group fully represents all types of RA patients and suggests that the results readily translate to clinical practice settings in which early RA patients will be treated to target, upon diagnosis. Moreover, projected disease activity outcomes from the modelled cohorts were shown to closely approximate real world outcomes of patients seen in practice, as displayed in figure 2. A limitation of this study is that due to a lack of data on productivity loss and employment for all patients, there was no possibility to collect out-of-pocket expenses or use the friction cost method, which should be considered when comparing our results to results obtained in other cost-effectiveness studies in this patient population.

This study analysed data from the Dutch societal perspective. Generalizing the results to other European countries should be handled carefully as medication list prices can vary across European countries, even in spite of the external reference pricing system that many EU members apply [88]. Additionally, the EQ-5D tariff, which is used to calculate the utility score from the EQ-5D questionnaire, varies per country. However, EULAR recommendations for T2T management apply to all European countries. Additionally, clinical features, like the DAS28 patterns, are not likely to vary per country. All in all, the model could be adapted to give an accurate representation of a different (European) country by adjusting the medication prices and the EQ-5D tariff.

In summary, the results of this study suggest that treating recently diagnosed RA-patients to the target of remission according to this strategy of initial combination therapy not only results in more patients in beneficial states of disease activity (remission or low disease activity) compared with initial combination therapy, but also at lower costs.

Supplemental Material

Supplement Table 1:

	RIC I (N = 509)	**RIC II (N= 381)**
Female, sex n (%)	321 (63.1%)	227 (59.6%)
Age, mean ± SD years	58.5 ± 14.5	59.8 ± 13.7
DAS28, mean ± SD	4.3[k] ± 1.5	4.1[c] ± 1.6
Number of TJC, median (IQR)	3.0 (1.0 – 7.0)	2.0 (0.0 – 6.0)
Number of SJC, median (IQR)	5.0 (2.0 – 9.0)	3.0 (1.0 – 8.0)
ESR (mm/h), median (IQR)	22.0[i] (11.0 – 38.0)	20.0 [a] (9.0 – 36.0)

CRP (mm/h), median (IQR)	10.0[j] (5.0 – 20.0)	7.0[b] (3.0 - 18.0)
Anti-CCP positive, n(%)	253[l] (58.2%)	133[d] (58.3%)
RF positive, n (%)	355[m] (71.9%)	145[e] (59.2%)
SF36-PCS, mean ± SD	36.8[n] ± 8.7	38.6[f] ± 9.6
SF36-MCS, mean ± SD	48.1[o] ± 11.7	46.5[f] ± 12.2
HAQ-SDI, median (IQR)	1.3[p] (0.6 – 1.9)	0.8[g] (0.3 – 1.4)
BMI kg/m², mean ± SD	26.6[q] ± 4.9	26.2[h] ± 4.5

[a] = 377; [b] = 362; [c] = 367; [d] = 228; [e] = 245; [f] = 249; [g] = 253; [h] = 124; [i] = 500; [j] = 486; [k] = 500; [l] = 435; [m] = 494; [n] = 401; [o] = 401; [p] = 478; [q] =383

DAS28 = Disease Activity Score in 28 joints; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; TJC = tender joint count; SJC = swollen joint count; HAQ-SDI = Health Assessment Questionnaire disability index (standard scoring); SF-36 = Short-Form 36 health survey (version 2); PCS = physical component summary; MCS = mental component summary; RF = rheumatoid factor; Anti-CCP = anti-cyclic citrullinated peptide; BMI = body mass index; RIC = Remission Induction Cohort

Chapter 3

Chapter 3 – Combining Data Sources in Cost-Effectiveness Analyses

Based on: Cost-effectiveness of a JAK1/JAK2 inhibitor vs. a biologic disease-modifying antirheumatic drug (bDMARD) in a Treat-to-Target strategy for rheumatoid arthritis

Van De Laar CJ, Oude Voshaar MA, Fakhouri WK, Zaremba-Pechmann L, De Leonardis F, De La Torre I, Van De Laar MA. Cost-effectiveness of a JAK1/JAK2 inhibitor vs a biologic disease-modifying antirheumatic drug (bDMARD) in a treat-to-target strategy for rheumatoid arthritis. ClinicoEconomics and Outcomes Research. 2020 Apr 15:213-22.

Abstract

Background

Baricitinib is a janus kinase (JAK1/JAK2) inhibitor developed for the treatment of patients suffering from rheumatoid arthritis (RA). Treating RA to the target of remission is current common practice. Cost-effectiveness of different Treat-to-Target (T2T) strategies, especially ones including new treatments is important for development and preference policy for treatment centres. European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) guidelines are currently unclear about preference between a JAK1/JAK2 versus a biological Disease-Modifying Anti Rheumatic Drug (bDMARD).

Objectives

The main goal of this paper was to evaluate the cost-effectiveness of baricitinib versus first biological for methotrexate inadequate responders in a T2T strategy using a Markov model that incorporates hospital costs as well as societal costs. Costs and utilities over five years were compared between the two strategies.

Methods

A Monte Carlo simulation model was developed to conduct cost-utility analysis from the societal perspective over 5 years. Health states were based on the DAS28-erythrocyte sedimentation rate (ESR) categories. Effectiveness of baricitinib was retrieved from randomized controlled trials. Effectiveness of all other treatments, health state utilities, medical costs, and productivity loss were retrieved from the Dutch RhEumatoid Arthritis Monitoring (DREAM) cohorts. Annual discount rates of 1.5% for utility and 4% for costs were used. Probabilistic sensitivity analysis was employed to incorporate uncertainty and assess robustness of the results.

Results

Probabilistic Sensitivity Analysis results showed the baricitinib strategy yielded lower costs and higher utility over a 5-year period. The estimated difference between the strategies in QALYs accrued over 5 years was small, while the difference in expected costs between the strategies was considerable, leading to a large ICER. Scenario analyses showed the baricitinib strategy to be cost-effective in both the moderate and severe RA populations.

Conclusions

In this Monte Carlo simulation, baricitinib strategy was cost-effective compared to DREAM T2T strategy in a T2T setting for RA patients with moderate and severe disease activity. Results suggest that the use of a JAK1/JAK2 inhibitor instead of a bDMARD in a T2T approach is cost-effective in csDMARD refractory RA patients.

Introduction

Rheumatoid arthritis (RA) is a systemic autoimmune disease with alternating periods of low, moderate and/or high disease activity, usually characterized by inflammation of the synovium. RA can be chronic in which case it may lead to functional impairment and disability if left untreated [54]. Accumulating evidence suggests that optimal treatment outcomes may be obtained with timely initiated tight control management of disease activity.

This approach to RA management, also referred to as Treat to Target (T2T), involves setting disease activity targets at the onset of treatment. Disease activity is then routinely measured using one of several available composite disease activity indices and medication adjustments are made contingent on the clinical composite score. If the disease activity target is not met, treatment is intensified (e.g. by increasing doses, combining treatments or switching to a different drug). If disease activity is on target during successive measurements, medications may be tapered [89].

All major international management guidelines currently endorse a T2T approach to RA management [58,90]. Within this general framework, the current recommended approach is to start patients on conventional synthetic disease-modifying antirheumatic drugs (csDMARDs). csDMARDs are a group of traditional RA medications. While csDMARDs have not been developed to target a molecular structure, they have become known for their disease modifying properties and are recommended first line agents in light of their favourable cost profile [91,92]. If the treatment target cannot be achieved using csDMARDs, international guidelines recommend the initiation of a biological DMARD (bDMARD) or a synthetic drug designed to target particular molecular structures. This class of targeted synthetic DMARDs (tsDMARDs) is relatively new in rheumatology and international recommendations place them at the same level as bDMARDs, with the csDMARD inadequate responders (IR) population as the earliest possible time point for use in the therapeutic algorithm. Few studies have yet directly compared these agents. Of note, the Janus kinase (JAK)1/JAK2-inhibitor baricitinib, now approved in more than 50 countries worldwide, was found to be statistically superior for American College of Rheumatology 20 criteria (ACR20) and disease activity score 28 - C-reactive protein (DAS28-CRP) mean change at week 12 against bDMARD adalimumab with background methotrexate (MTX), when administrated in combination with MTX in a population of moderate-to-severe RA patients who are MTX- IR [93]. Currently, there is little information on the long-term cost-effectiveness of JAK-inhibitors versus bDMARDs as part of contemporary T2T based management strategies.

The aim of this study was to simulate long-term patient outcomes and cost-effectiveness of two treatment strategies in which either a JAK1/JAK2-inhibitor (baricitinib) or a bDMARD is initiated in csDMARD-IR by developing a Monte Carlo simulation Markov model.

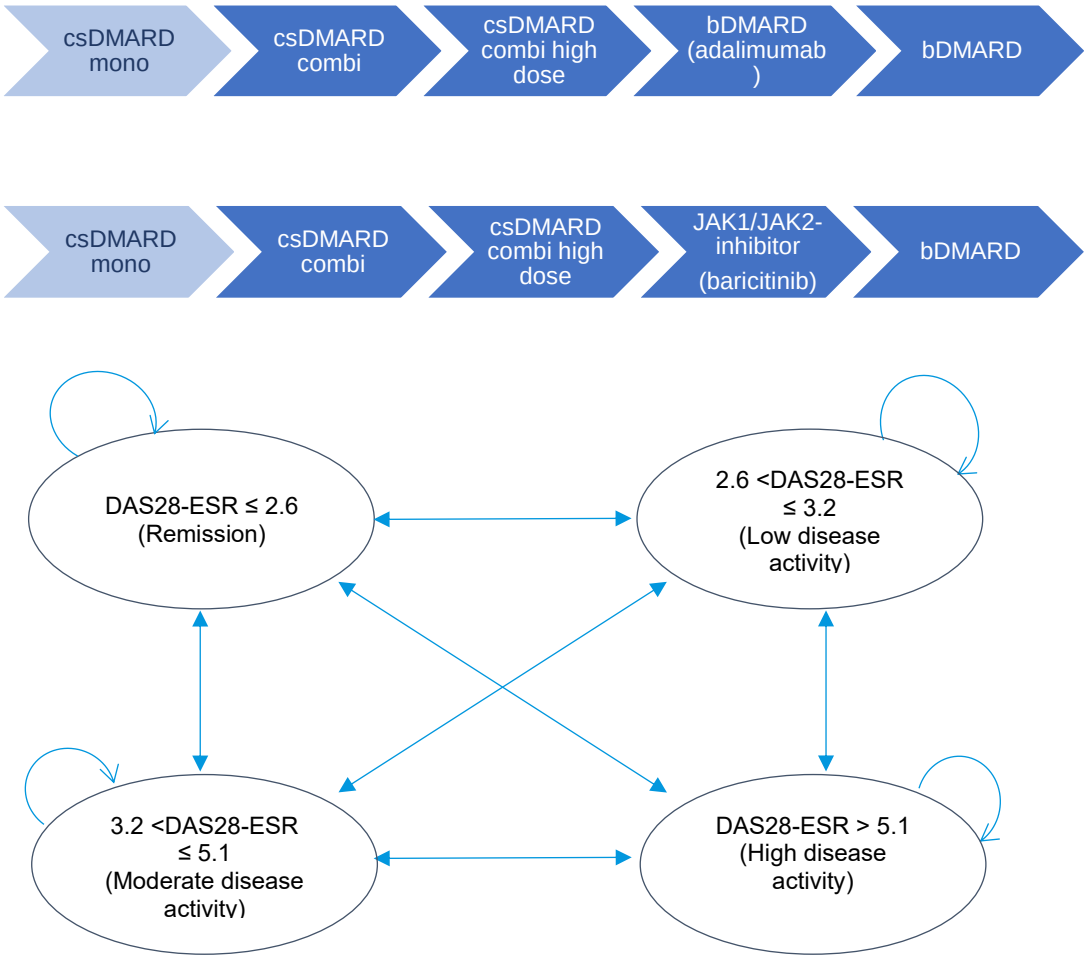
Materials and Methods

A cost utility analysis (CUA) was conducted from the Dutch societal perspective using an individual-sampling Monte Carlo simulation Markov model. The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist was followed in reporting of the analyses. The model was subjected to checks and validation by an external party (IQVIA).

Analytic Framework and treatment strategies

The model was developed using TreeAge Pro 2018, R2 (TreeAge Software, Williamstown, MA) to simulate changes in disease activity, health utility, and costs in patients with early RA (within one year after diagnosis) receiving treatment based on the T2T principle as applied in the Dutch RhEumatoid Arthritis Monitoring registry (DREAM) cohorts from the moment of diagnosis onwards [64,94]. In the model, patients transition between disease activity states on a 12-weekly basis. Four commonly distinguished clinically relevant disease activity states based on the Disease Activity Score 28-erythrocyte sedimentation rate (DAS28-ESR) categories, are used in the model: Remission (DAS28-ESR≤2.6), low disease activity (2.6<DAS28-ESR≤3.2), moderate disease activity (3.2<DAS28-ESR≤5.1), and high disease activity (DAS28-ESR>5.1) [95]. Patients enter the model with an assigned disease activity state. Patient distribution over health states at baseline is based on the observed distribution in the DREAM cohorts. Patients can move from any state to any of the others states every cycle. See Figure 1 for a graphic display of all possible transitions. The transition probabilities depend on the current health state and the treatment that is currently used. This model structure was adapted from a model previously introduced by Welsing et al. [74,96] and extended by Schipper et al.[73]. Welsing et al. have demonstrated that model-based disease activity trajectories over a period of 5 years closely approximate the disease activity course actually observed in the cohort of patients that transition probabilities were derived from.

Figure 1: Treatment Strategies, Model Health States, and possible transitions



csDMARD = conventional synthetic disease-modifying antirheumatic drug, mono = monotherapy, combi = combination therapy, bDMARD = biological disease-modifying antirheumatic drug, JAK1/JAK2-inhibitor = Janus Kinase 1, 2, inhibitor, DAS28-ESR = Disease Activity Score 28-erythrocyte sedimentation rate

In alignment with current management recommendations and the T2T principle, all treatment changes in the model in the present study are governed by the disease activity state of the patient. For all strategies, treatment was targeted at achieving DAS28-ESR remission [64,94]. If a patient has been in non-remission for 1 cycle (12 weeks) the next treatment is selected. Patients that are in remission and remain there for 2 cycles (24 weeks) move back to previous medication. This so-called tapering of medication after two cycles is in accordance with the current rheumatology guidelines [58,90]. Eventually, in case of long-term sustained remission, patients can reach medication-free remission through tapering. Patients that go through tapering will return to their last effective medication in case of a flare (DAS28-ESR>2.6). Due to reimbursement restrictions in the Netherlands, the first bDMARD will only be selected if DAS28-ESR>3.2, i.e., moderate or high disease activity.

Two treatment strategies were evaluated and compared. In the first modelled strategy (DREAM T2T), patients are initially treated with a low dose csDMARD combination therapy. In case remission is not achieved, higher dose csDMARD combination is next, before a bDMARD (adalimumab), and ultimately a (second) bDMARD are initiated. Adalimumab was chosen as first bDMARD after inadequate response to TNF-α blockers due to its common use in the Netherlands and internationally. This is also reflected in data availability for effectiveness of TNF-α blockers. In the Baricitinib Strategy, the first bDMARD (adalimumab) is replaced by baricitinib. When patients experience sustained remission (two consecutive cycles spent in remission) their medication is tapered. In case of sustained remission, a patient moves back one treatment step. When the patient is in sustained remission again, he/she can move another step, until arriving at medication-free remission. There is an additional 'taper step' between each strategy's initial treatment and medication-free remission, this is MTX monotherapy in both strategies. For example, a patient in the DREAM T2T strategy on adalimumab who is in sustained remission moves back to MTX Combination therapy, in case of another sustained remission, moves to MTX Monotherapy (the additional taper step before medication-free remission), then in case of another sustained remission will go to medication-free remission. In case a flare happens at any point during the tapering process, the patient returns to the last effective treatment.

The difference between these strategies lies in the initiation of baricitinib vs. a bDMARD in csDMARD-IR population. Patients enter the model upon diagnosis and are distributed over the four DAS28 health states as observed in the patient cohorts used as data input. See Figure 1 for an outline of the exact treatment strategies and health states used in the model.

Analyses

Expected values for costs and utility over five years were obtained and uncertainty around these estimates was quantified using Probabilistic Sensitivity Analysis (PSA) with Monte Carlo simulation using 1500x150 samples. For each input parameter, several potentially appropriate distributions according to literature recommendations [78] were selected and compared their relative fit using Chi-square tests. Costs and utility were discounted annually at 4% and 1.5%, respectively, according to the Dutch Costing Guidelines for Economic Evaluation [71].

Additionally, scenario analyses were performed to explore what the effect would be of lowered drug prices. Specifically, analyses were run where lowered drug prices for adalimumab and baricitinib were put into the model. It can be supposed that due to the introduction of biosimilars, prices of expensive drugs and specifically adalimumab could be lower than the list prices. To take possible discounts or other agreements into consideration the favourability of the two strategies (DREAM T2T strategy and baricitinib 2nd strategy) will be assessed at different levels of lowered medication prices.

Model inputs

All model inputs can be found in Tables 1, 2, and 3.

Efficacy

For each treatment in the model, a transition matrix, **P**, was used as input. There are 7 treatments in each strategy. These matrices describe the probabilities that patients move from DAS28 based health state *i (rows)* to health state *j (columns),* with row and column labels: 1 = Remission, 2 = Low, 3 = Moderate, and 4 = High disease activity. Transition probabilities are assumed to be independent of time and transitions are assumed to occur in evenly spaced time intervals of three months, so that if Y_t is the health state at a particular time point *t*:

$$\mathbf{P}_{ij} = P(Y_t = j | Y_{t-3\,months} = i)$$

and the three-monthly transition probabilities can be estimated simply from the observed sample proportions as follows:

$$\widehat{\boldsymbol{P}}_{ij} = \frac{n_{ij}}{\sum_{k=1}^{m} n_{ik}}$$

where n_{ij} is the number of times the transition from *i* to *j* was observed in the total dataset and the denominator is the total number transitions out of *i,* with *m* = the number of DAS28 states (=4). The model will first select the correct matrix (dependent on current treatment), and then find the probabilities of moving to any of the four health states for each sampled patient. An example of such a transition matrix is provided below, with transition probabilities for MTX monotherapy derived from DREAM observational data. The transition probabilities for all treatments other than baricitinib were calculated using the DREAM observational data. Approximately 4000 observations were used. For baricitinib, a matrix of three-monthly transition probabilities was calculated from data provided by Lilly from the BEAM, BUILD, and BEGIN studies.

Transition probabilities for each treatment step that does not involve baricitinib were estimated using data from the ongoing "Remission Induction in Early Rheumatoid Arthritis" (Netherlands Trial Register: NTR578) cohorts in the DREAM registry [64,94]. Each matrix of observed transitions was filled using data of the first year of follow up (i.e. a maximum of four (~12 weekly ±~6 weeks) visits) of patients' first and second therapies with that specific treatment. Transition probabilities were assumed to be independent of time and to occur in evenly spaced time intervals of twelve weeks. The assumption of time independence was tested using Chi-square tests directed at deviations of individual empirical matrices (e.g. the matrix of observed transitions from baseline to week 12) from the corresponding input matrix, which was obtained by pooling all observed transitions. The model selects the correct matrix (dependent on current treatment) and then finds the probability of moving to any of the four health states (dependent on current health state) for each sampled patient. The EuroQol-5 dimensions (EQ-5D-3L) [69] questionnaire was used to measure the quality of life in all four respective health states and valued using the Dutch tariff [70].

Baricitinib transition probabilities

Real-world effectiveness data on baricitinib in RA patients has not yet been collected long enough to use in this model, at DREAM or any other registries. Transition probabilities for baricitinib were therefore derived from pivotal trial data [93,97]. Transition probabilities for baricitinib were adapted by assuming that for each element in the matrix of estimated baricitinib transitions the

relative risk (RR) of observing the transition for baricitinib compared with an available common comparator would be the same in Lilly baricitinib pivotal trials and the DREAM T2T setting. Supplemental Material provides a more in-depth explanation of this methodology.

Table 1: Model Input Distributions

Markov States	Utility (EQ-5D)			Healthcare costs (€)			Sick days		
	Mean (SD)	Distribution	Parameters	Mean (SD)	Distribution	Parameters	Mean (SD)	Distribution	Parameters
Remission (n=1272)	0.76 (0.16)	Normal	μ = 0.76 σ = 0.16	€198.34 (317.54)	Gamma	α = 0.38 β = 508.91	0.34 (1.71)	Normal	μ = 0.34 σ = 1.71
Low disease activity (n= 653)	0.71 (0.20)	Normal	μ = 0.72 σ = 0.19	€286.16 (486.89)	Gamma	α = 0.34 β = 830.47	0.35 (1.99)	Gamma	α = 0.03 β = 11.51
Moderate disease activity (n = 1382)	0.64 (0.22)	Normal	μ = 0.64 σ =0.22	€360.41 (548.12)	Normal	μ = 360.41 σ = 548.12	0.38 (2.92)	Gamma	α = 0.04 β = 9.47
High disease activity (n=408)	0.52 (0.27)	Normal	μ = 0.52 σ = 0.27	€ 475.63 (683.66)	Gamma	α = 0.47 β = 994.66	0.70 (2.57)	Gamma	α = 0.07 β = 9.56

The fit of distributions was compared using the fit statistics provided by "Easyfit". Best fitting distributions were selected based on Kolmogorov Smirnov statistics, Chi-Squared statistics, and the Anderson Darling statistic.

Treatment cost

Table 2 Medication Costs

Drug	Dosage	Price per 12 weeks (€)
MTX monotherapy	20mg once a week	72.23
csDMARD combination low dosage	MTX 20mg once a week, HCQ 400mg once daily	95.15
csDMARD combination high dosage	MTX 25 to 30[1] mg once a week, HCQ 400mg once daily	127.19
Baricitinib	4mg once daily	3,307.40
Adalimumab	40mg once a week	3,479.85
Biological (etanercept, rituximab, infliximab, abatacept, or tocilizumab)	etanercept: 50mg once a week, rituximab: IV twice 1000mg, infliximab: 3mg/kg body weight, abatacept: <60 kg: 500 mg; 60–100 kg: 750 mg; >100 kg: 1000 mg, every 3 weeks. tocilizumab: 8mg/kg body weight, every 4 weeks.	3,819.50

MTX = Methotrexate, csDMARD = conventional synthetic disease-modifying antirheumatic drug, HCQ = Hydroxychloroquine

Prices are derived from the Dutch Pharmacy Purchase price list.

[1]Price calculated according to observed distribution in DREAM cohorts over 25mg and 30mg MTX prescribed.

Medication costs were calculated using the Dutch Pharmacy Purchase prices using the standard Dutch dosage. MTX/Hydroxychloroquine (HCQ) low dosage is daily 20mg MTX and twice-daily HCQ 200mg. MTX/HCQ high dosage refers to daily 25 or 30mg MTX (where 80% is on 30mg and 20% on 25, as observed in DREAM Remission Induction Cohort) and twice-daily HCQ 200mg. Adalimumab has a dosage of 40mg every other week. Baricitinib has a dosage of 4mg once daily. Next bDMARD consists of either etanercept (50mg per week), rituximab (1000mg every six weeks), infliximab (300mg every 4 weeks), abatacept (750mg every 3 weeks), and tocilizumab (800mg every 4 weeks) with equal probabilities.

For the sensitivity analysis focusing on lowered prices for the expensive medication, adalimumab and baricitinib, multiple price levels were considered. List price, 75% of list price, 50% of list price and 25% of the list price.

Other costs

Table 3: Other Costs per DAS28-ESR Disease Activity State

Variable	DAS28-ESR Remission	Low DAS28-ESR	Moderate DAS28-ESR	High DAS28-ESR
Healthcare costs (€)	198.34	286.16	360.41	475.63
Sick days cost (€)	78.20	80.50	87.40	161.00
Utility (QALY)	0.76	0.71	0.64	0.52

DAS28-ESR = Disease Activity Score 28-erythrocyte sedimentation rate, QALY = Quality-Adjusted Life Year
Costs were retrieved from www.medicijnkosten.nl, the Dutch list prices.

Other costs included in the model are Healthcare Consumption Costs and work productivity loss due to sick leave. Both can be found in table 3. Both cost sources were collected in RA patients in the Netherlands, being treated in a T2T setting. Healthcare resource usage was assessed using a patient questionnaire and costs were calculated based on the amounts of resource use and unit costs [71]. The following information on healthcare resource usage is routinely collected in DREAM by patient questionnaire: number of diagnostic tests; hospital admission days; specialist (rheumatologist and other), nurse and general practitioner visits; visits to psychologists, psychiatrist, and physical therapist, and hours of formal and informal care. These were valued using the Dutch Costing guidelines published by the Dutch Health Care Insurance Board [71]. Data were also collected on productivity loss due to sick leave using the Human Capital method (includes hourly salary, social security fees, private car use, administration costs). Sick leave cost was also included in the model, as this means a loss of productivity and costs for the society. Sick leave cost was valued per diem [72], multiplied with the observed percentage of patients in the workforce (in DREAM registry) and added to the total costs per health state.

Results

Tables 1 and 2 display values and distributions for all parameters. In the PSA, the baricitinib strategy dominated the DREAM T2T strategy, with lower expected costs, and higher expected utility over five years (see Table 4). The Incremental Cost-Effectiveness Ratio (ICER) was estimated to be € -238,418. The negative ICER in this case reflects the dominance of baricitinib strategy, it accumulates lower cost and higher Quality-Adjusted Life Years (QALYs) over 5 years. The estimated difference between the strategies in QALYs accrued over 5 years was small, while the difference in expected costs between the strategies was considerable, leading to a large ICER.

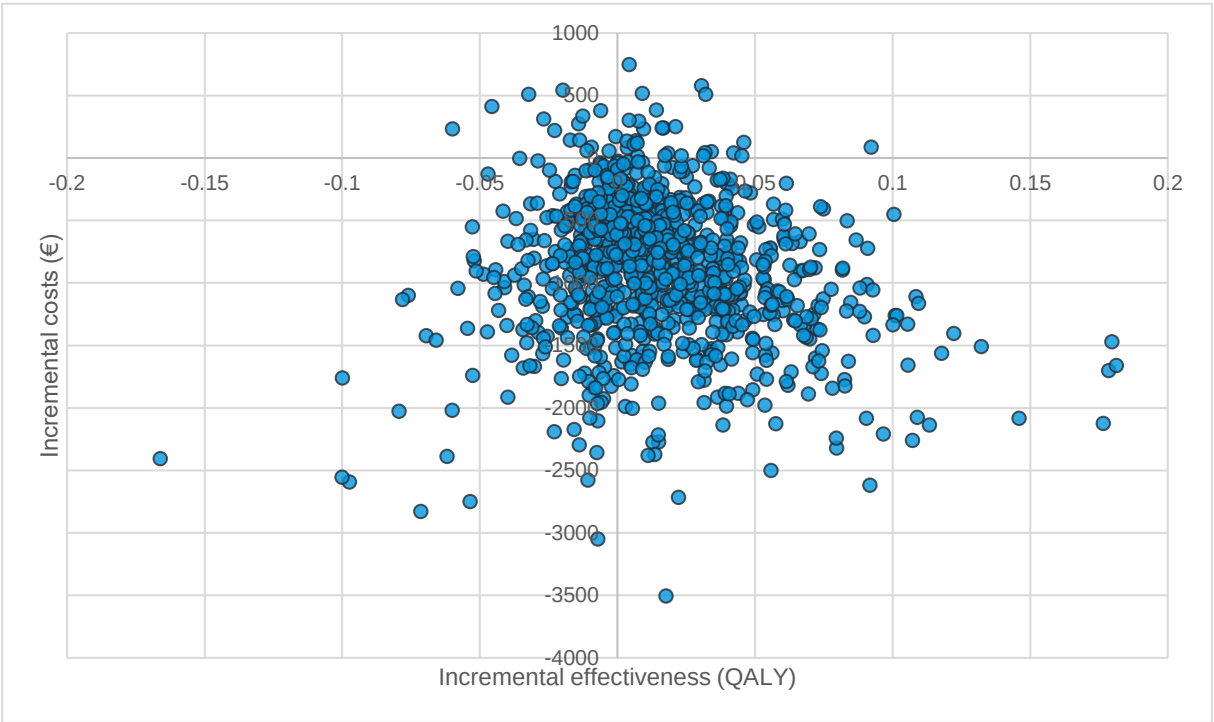
Table 4: Probabilistic Sensitivity Analysis Results

	DREAM T2T Strategy	Baricitinib Strategy
Costs € (95% CI)	14288.36 (10104.68-19053.39)	13430.57 (9475.103-17848.92)
QALYs (95% CI)	3.5607 (2.4028-4.7236)	3.5643 (2.4042-4.7315)
ICER	-	Dominant (-238,418)
NMB (at WTP = €60,000)	199,354.9	200,428.6

CI = confidence intervals, QALYs = Quality Adjusted Life Years, ICER = Incremental Cost-effectiveness Ratio, NMB = Net Monetary Benefits, WTP = Willingness to Pay

Figure 2 shows a plot of the distribution of ICER estimates resulting from the Monte Carlo iterations over the cost-effectiveness plane. Almost all (96.07%) estimates of incremental costs were negative (i.e. baricitinib strategy incurred less costs); while in 64% of the model replications, the ICER fell in the southeast quadrant (i.e. baricitinib incurred less costs and provides better health outcomes). As can be seen in table 5, baricitinib was likely to be cost effective at the willingness to pay (WTP) threshold of €60,000 with 81.94% of all iterations below this threshold. 64.00% of iterations were cost saving. Even at more extreme WTP thresholds of €100,000 and €500,000, baricitinib strategy was cost-effective in 75% and 55% of iterations, respectively. Scenario analyses showed the baricitinib strategy to be cost-effective in both the moderate and severe RA populations.

Figure 2: ICER PSA Results Baricitinib Strategy vs. DREAM T2T Strategy



QALY = Quality-Adjusted Life Year
500 iterations displayed

Table 5: ICER Report

ICER Quadrant	Below WTP threshold (cost-effective)	Above WTP threshold	Total proportion
North East (IE>0, IC>0)	2.47%	0.20%	2.67%
North west (IE<0, IC>0)	0.00%	1.27%	1.27%
South west (IE<0, IC<0)	15.47%	16.60%	32.07%
South east (IE>0, IC<0)	64.00%	0.00%	64.00%
Total:	81.94%	18.07%	100.00%

WTP = Willingness to Pay, IE = Incremental effect, IC = Incremental costs, ICER = Incremental Cost-effectiveness Ratio. Percentages based on a WTP-threshold of €60.000

Scenario analyses

Disease activity at baseline

Analyses in which the baseline distribution of patients entering the model were altered have also been run. The base case analysis used a patient distribution over DAS28-ESR states as observed in the DREAM RIC. The analyses below used all patients having a high DAS28-ESR (DAS28-ESR>5.1) at baseline and a second analysis in which all patients have moderate DAS28-ESR (3.2<DAS28-ESR≤5.1) at baseline.

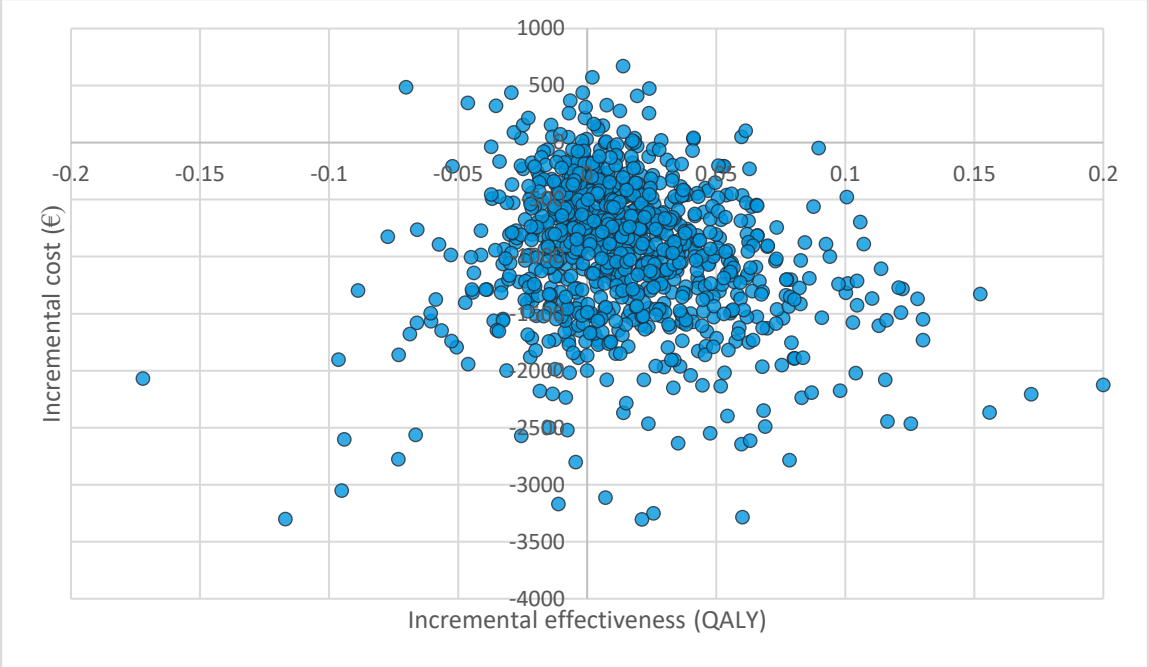
The PSA results show that the baseline distribution does not alter the conclusion from the base case analysis. Bari second remains a dominant strategy in both scenario's outlined in table 6. Mean costs over five years in Bari second are around €1,000 less expensive than in DREAM T2T, and mean QALYs over five years are slightly higher. This results in a negative ICER for Bari second as compared to DREAM T2T. Bari first is not cost-effective in the high DAS28 scenario, and an inferior strategy in the moderate DAS28 scenario. NMBs are highest for Bari second in both scenario's.

Table 6: Sensitivity Analysis: PSA Results with different baseline DAS28-ESR distribution

High DAS28 at baseline	DREAM T2T	Bari first	Bari second
Costs (€)	15078.77 (10154.53-20515.62)	52884.5 (48756.28-56744.11)	14169.29 (9606.301-19136.31)
Utility (QALYs)	3.5571 (2.4083-4.7304)	3.5641 (2.4141-4.7187)	3.5608 (2.4079-4.7396)
ICER	-	5,356,777	-243,631
NMB (at WTP = 60,000)	198346.1	160963.8	199479.5
Moderate DAS28 at baseline			
Costs (€)	14500.04 (10108.83-19509.37)	53077.98 (48926.22-56969.31)	13628.43 (9487.835-18428.26)
Utility (QALYs)	3.5599 (2.4046-4.7179)	3.5598 (2.4054-4.7133)	3.5636 (2.4054-4.7278)
ICER		-2.6E+08	-238685
NMB (at WTP = 60,000)	199095.9	160508.9	200186.6

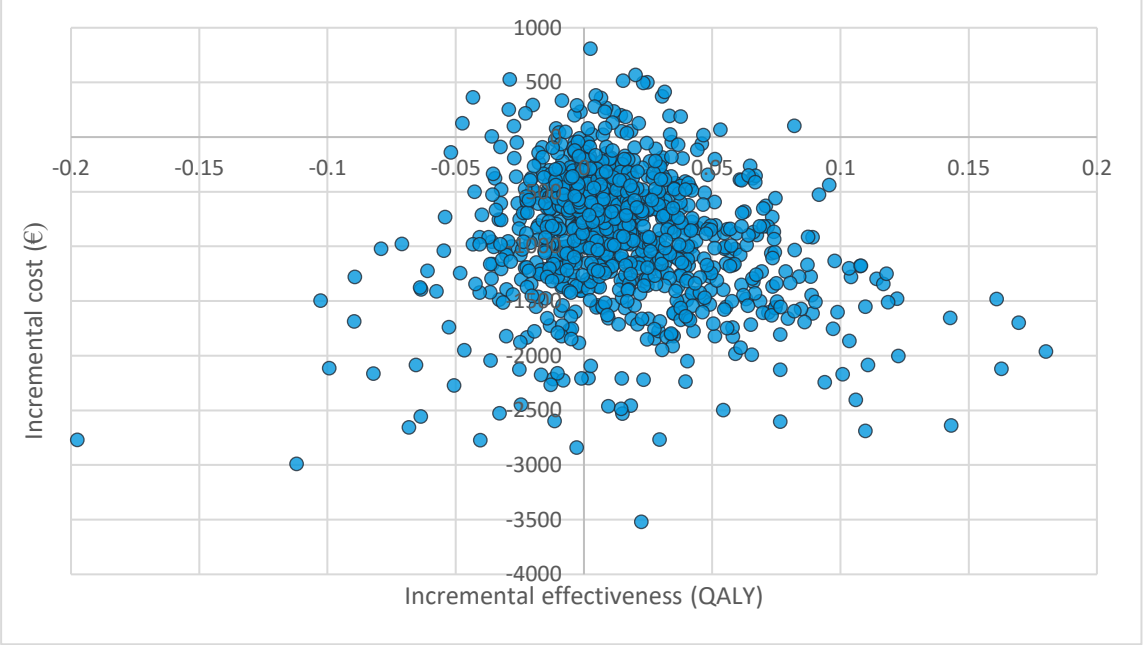
WTP = Willingness to Pay, IE = Incremental effect, IC = Incremental costs, ICER = Incremental Cost-effectiveness Ratio, NMB = Net Monetary Benefits

Figure 4: Sensitivity Analysis: ICER plane with all patients high DAS28 at baseline



QALY: Quality-Adjusted Life Years

Figure 5: Sensitivity Analysis: ICER plane with all patients moderate DAS28 at baseline



QALY: Quality-Adjusted Life Years

Figure 4 displays the ICER plane of Bari second vs. DREAM T2T for High DAS28 at baseline scenario. It looks very similar to the base case scenario. This scatterplot shows the incremental cost and incremental effectiveness and displays how a large proportion (63.40%) of iterations is in the southeast quadrant: which means they have positive incremental effects and negative incremental

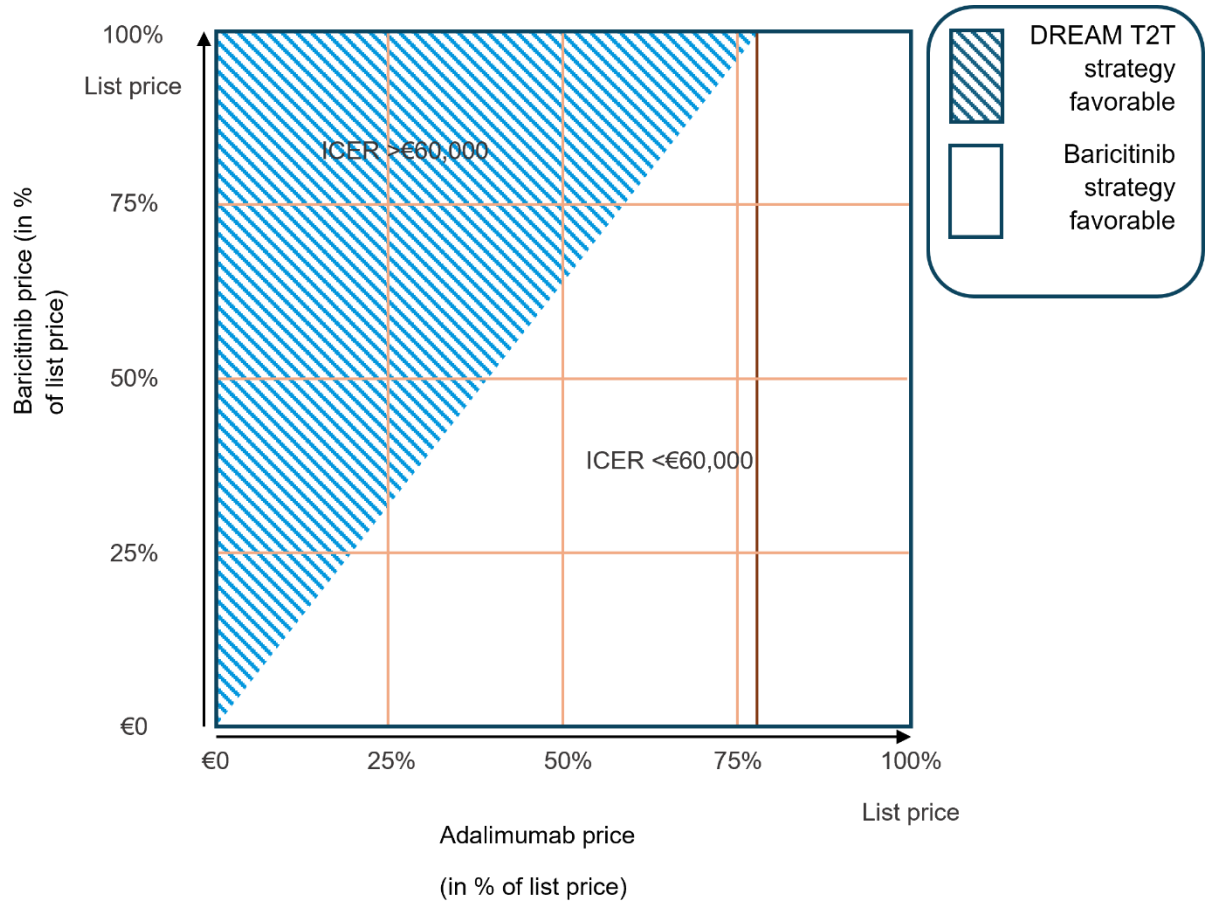
costs and is thus cost-saving. This means that in 64.00% of cases, implementing the Bari second strategy instead of DREAM T2T strategy will cost less while it generates more QALYs. An even larger proportion of all iterations is cost-effective (at WTP of €60,000), meaning that they cost less than €60,000 for one additional QALY when comparing Bari Second to DREAM T2T. 80.80% of all iterations is cost-effective at a WTP threshold of €60,000.

Figure 5 displays the ICER plane of Bari second vs. DREAM T2T for the moderate DAS28 at baseline scenario. This scatterplot shows the incremental cost and incremental effectiveness and displays how a large proportion (64.33%) of iterations is in the southeast quadrant: which means they have positive incremental effects and negative incremental costs and is thus cost-saving. This means that in 64.00% of cases, implementing the Bari second strategy instead of DREAM T2T strategy will cost less while it generates more QALYs. An even larger proportion of all iterations is cost-effective (at WTP of €60,000), meaning that they cost less than €60,000 for one additional QALY when comparing Bari Second to DREAM T2T. 81.40% of all iterations is cost-effective at a WTP threshold of €60,000. In the moderate DAS28 at baseline scenario, a slightly larger proportion of iterations is cost-saving, but all three scenarios yield very similar results.

Price variation

Additionally, a scenario analysis was run in which lowered prices were considered for adalimumab and baricitinib. Figure 6 shows the accumulated results. For both adalimumab and baricitinib various price levels were considered: list price, 75% of list price, 50% of list price and 25% of list price. Figure 6 shows whether the ICER is below or above the WTP threshold of €60,000, and thus whether the DREAM T2T strategy or the Baricitinib strategy is favourable.

Figure 6: Overview of favourability of baricitinib 2nd strategy vs. DREAM T2T strategy at different drug price levels for baricitinib and adalimumab



This figure shows the favourability of the two strategies, DREAM T2T strategy versus Baricitinib strategy. A WTP-threshold of €60,000 was used to assess the ICERS at each combination of prices for baricitinib and adalimumab. ICER = Incremental Cost-Effectiveness Ratio. WTP-threshold = Willingness-to-Pay threshold.

Discussion

In this study, the cost-effectiveness of a T2T strategy in which JAK1/JAK2-inhibitor baricitinib is tried in csDMARD-IR relative to a T2T strategy in which a bDMARD is tried after csDMARD therapy was evaluated. The results suggest that treatment with baricitinib in csDMARD-IR, as part of a tight-control, step up approach is very likely to incur lower costs and likely to be cost effective at various cost-effectiveness thresholds. However, in case of biological and biosimilar drugs, real-world prices might differ from the commonly used list prices. The results of the scenario analysis suggest that differences in pricing could have an effect on cost-effectiveness profiles of different RA treatment strategies.

The ACR and EULAR management guidelines currently do not distinguish between the relative merit of JAK-inhibitors such as baricitinib compared with bDMARDs as components in contemporary T2T based management strategies, in light of the limited real-world evidence on their relative effectiveness. However, these recommendations are solely based on direct

comparisons of tsDMARDs and bDMARDs in clinical trials. While this provides relevant data on the relative safety and efficacy, integrating health-economics considerations play a crucial factor in the decision-making process for drug prescriptions in RA. In the past, health economic considerations have helped shape T2T sequences. The present study is the first to show the value of JAK1/JAK2-inhibitor baricitinib relative to the current second line treatment from a societal perspective.

The model used in this study was initially developed and validated by Welsing et al, and later adapted by Schippers et al to model sequential treatment outcomes in RA patients. Unlike many previously published model-based health economic evaluations, the status of patients does not deteriorate over time in this model. This is consistent with studies describing 5-10-year outcomes of early RA T2T cohort studies and was therefore considered appropriate for our current study.

A strength of this work is that, whenever possible, the DREAM daily clinical practice data as inputs for the model were used. Most modelling studies in RA rely on clinical trial data inputs.[98] Typically, patients included in clinical trials are younger, have more severe disease, and fewer comorbidities compared with patients seen in daily practice.[61,99] The use of clinical trial data in health-economic modelling studies therefore limits the generalizability of the results and might lead to overestimation of health benefits, since these characteristics have previously been found to be associated with achieving treatment response [100]. However, unfortunately, real-world effectiveness data on baricitinib in RA patients is not yet available. Therefore, estimates of baricitinib transition probabilities were obtained by assuming that the relative effectiveness of baricitinib and bDMARD would be the same in daily clinical practice and the clinical trial setting. This model thus incorporates data from both the DREAM daily clinical practice data and baricitinib clinical trial program to address the research question. Other limitations of this study are that no individual patient expenses were available in the DREAM cohort study, due to which there were no out-of-pocket costs included in this study. This does not represent a limitation for internal comparison of the two strategies; however, external comparisons should be conducted carefully. Drug discontinuation is not incorporated in the model, however, in the DREAM cohort it was shown that adherence to the protocol was good [65].

Conclusions

In this Monte Carlo simulation, baricitinib strategy was cost-effective compared to DREAM T2T strategy in a T2T setting for RA patients with moderate and severe disease activity. Scenario analysis showed similar results for moderate and severe populations analysed separately. Future analyses using real-world effectiveness data is important to validate these results.

Supplemental material

Baricitinib transition matrices adaptation – Relative Risk Methods

Baricitinib has not yet been administered in the DREAM cohort, nor are there other data sources which could provide daily clinical practice data on baricitinib effectiveness. Transition probabilities for baricitinib were estimated by assuming that for each element in the matrix of estimated baricitinib transitions the relative risk (RR) of observing the transition for baricitinib compared with an available common comparator would be the same in Lilly baricitinib pivotal trials and the DREAM T2T setting. That is, if $RR_{L_{ij}}$ denotes the risk of observing the transition from disease activity state *i* to disease activity state *j* with baricitinib relative to the common comparator in the Lilly pivotal trial data, and $RR_{D_{ij}}$ denotes the relative risk of observing the same transition in the DREAM T2T data, then

$$\frac{RR_{L_{ij}}}{RR_{D_{ij}}} \equiv 1$$

for each combination of *i* and *j*. Note that this still allows the transition probabilities for the common comparator to differ between data sources, reflecting different data source specific prognostic factors.

Each $RR_{L_{ij}}$ could be readily calculated from the available data, since both baricitinib and the common comparator were administered in the pivotal trials i.e.

$$RR_{L_{ij}} = \frac{Lb_{ij}}{Lc_{ij}}$$

where *Lb* is the pivotal trial baricitinib matrix, and *Lc* is the pivotal trial common comparator matrix. In the DREAM T2T cohort, only transition probabilities for the common comparator are available. However, by assuming that

$$\frac{RR_{L_{ij}}}{RR_{D_{ij}}} = 1$$

initial estimates of the DREAM baricitinib transition probabilities could be obtained by multiplying the individual DREAM transition probabilities with the corresponding relative risks as follows:

$$pb_{ij} = Dc_{ij}\, RR_{D_{ij}}$$

where pb_{ij} is an initial estimate of a DREAM baricitinib transition probability and Dc is the DREAM common comparator transition matrix.

To ensure that the marginal row probabilities stay between 0 and 1 the estimated transition probabilities were rescaled by dividing each initial transition probability by the sum of the respective row probabilities to obtain “rescaled” transition probabilities:

$$Db_{ij} = \frac{pb_{ij}\, RR_{L_{ij}}}{\Psi_i}$$

where Db is a matrix of rescaled DREAM baricitinib transition probabilities, and Ψ_i is the sum of the possible numerators for row i. All calculations are provided in the supplemental data. * - upon request? - ***** The rescaled transition probabilities were in most cases only slightly different from the initial transition probabilities for the transitions from remission, mild and moderate disease activity to other disease states. However larger differences were sometimes observed for transitions from high disease activity to other disease activity states. This was most likely the result of data sparsity. Only ~3% of the observed transitions in the DREAM data were from a state of high disease activity, so that individual transitions sometimes had to be estimated with n < 10. Table 2 presents the "initial" estimates of DREAM baricitinib 4mg transition probabilities derived from RA-BEAM. Table 3 presents the corresponding "rescaled" estimates.

Supplement Table 1: Example matrix of "initial" estimates of RA-BEAM 4mg Baricitinib monotherapy transition probabilities

MTX mono	Remission	Low	Moderate	High
Remission	81.98%	11.76%	7.02%	0.05%
Low	60.74%	23.44%	19.56%	0.00%
Moderate	34.07%	21.32%	37.05%	5.91%
High	13.81%	33.28%	47.39%	19.54%

Supplement Table 2: Example of matrix of "Rescaled" estimates of RA-BEAM 4mg Baricitinib monotherapy transition probabilities

MTX mono	Remission	Low	Moderate	High
Remission	81.32%	11.66%	6.96%	0.05%
Low	58.55%	22.60%	18.86%	0.00%
Moderate	34.64%	21.68%	37.67%	6.01%
High	12.11%	29.19%	41.57%	17.14%

Transition matrices for the baricitinib first strategy were derived using RA-BEGIN, and the transition probabilities for the baricitinib replaces adalimumab strategy were derived using RA-BEAM and RA-BUILD data. Separate matrices were first obtained for both data sources separately and then combined by taking their unweighted mean.

Instead of the relative risk, other risk measures, such as the odds ratio or risk difference could have also been used in these calculations. Risk ratios were chosen because it would prevent unlikely transitions from dropping below 0%. However slightly different results might have been obtained if a different risk measure would have been used.

Transition Matrices adapted

Supplement Table 3

DREAM matrices

Adalimumab	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.81	0.12	0.07	0.00
Low Disease Activity	0.57	0.24	0.20	0.00
Moderate Disease Activity	0.34	0.24	0.36	0.07
High Disease Activity	0.14	0.20	0.43	0.23

Methotrexate	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.83	0.11	0.07	0.00
Low Disease Activity	0.59	0.24	0.18	0.00
Moderate Disease Activity	0.30	0.25	0.36	0.08
High Disease Activity	0.13	0.20	0.43	0.25

BEAM Matrices

Adalimumab	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.71	0.19	0.09	0.01
Low Disease Activity	0.28	0.37	0.34	0.02
Moderate Disease Activity	0.11	0.16	0.65	0.08
High Disease Activity	0.06	0.05	0.39	0.50

Baricitinib 4mg	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.72	0.19	0.09	0.00
Low Disease Activity	0.30	0.36	0.33	0.00
Moderate Disease Activity	0.11	0.15	0.67	0.07
High Disease Activity	0.06	0.09	0.43	0.43

Build Matrices

Adalimumab	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.59	0.26	0.15	0.00
Low Disease Activity	0.39	0.34	0.25	0.02
Moderate Disease Activity	0.10	0.12	0.66	0.13
High Disease Activity	0.00	0.01	0.27	0.71

Baricitinib 4mg	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.62	0.29	0.05	0.05
Low Disease Activity	0.31	0.34	0.31	0.03
Moderate Disease Activity	0.15	0.16	0.60	0.09
High Disease Activity	0.06	0.10	0.47	0.37

Estimated Matrices				
Estimated Baricitinib (initial, BEAM)	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.82	0.12	0.07	0.00
Low Disease Activity	0.61	0.23	0.20	0.00
Moderate Disease Activity	0.34	0.21	0.37	0.06
High Disease Activity	0.14	0.33	0.47	0.20
Estimated Baricitinib (Rescaled, BEAM)	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.81	0.12	0.07	0.00
Low Disease Activity	0.59	0.23	0.19	0.00
Moderate Disease Activity	0.35	0.22	0.38	0.06
High Disease Activity	0.12	0.29	0.42	0.17
Estimated Baricitinib (Initial, BUILD)	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.87	0.12	0.02	0.00
Low Disease Activity	0.46	0.24	0.23	0.00
Moderate Disease Activity	0.45	0.35	0.33	0.05
High Disease Activity	1.70	1.70	0.74	0.13
Estimated Baricitinib (Rescaled, BUILD)	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.86	0.12	0.02	0.00
Low Disease Activity	0.50	0.26	0.24	0.00
Moderate Disease Activity	0.38	0.30	0.28	0.05
High Disease Activity	0.40	0.40	0.17	0.03
Estimated Baricitinib (combined)	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.84	0.12	0.05	0.00
Low Disease Activity	0.54	0.24	0.22	0.00
Moderate Disease Activity	0.36	0.26	0.33	0.05
High Disease Activity	0.26	0.35	0.29	0.10
Estimated DREAM baricitinib 4mg	Remission	Low Disease Activity	Moderate Disease Activity	High Disease Activity
Remission	0.08	0.12	0.05	0.00
Low Disease Activity	0.54	0.24	0.22	0.00
Moderate Disease Activity	0.36	0.26	0.33	0.05
High Disease Activity	0.26	0.35	0.29	0.10

Chapter 4

Chapter 4 – Health Economic Modelling in Gout: Defining Health States by Pain Level

Published as: Model-based cost-effectiveness analyses comparing combinations of urate lowering therapy and anti-inflammatory treatment in gout patients

Van de Laar CJ, Janssen CA, Janssen M, Oude Voshaar MA, Al MJ, van de Laar MA. Model-based cost-effectiveness analyses comparing combinations of urate lowering therapy and anti-inflammatory treatment in gout patients. PloS one. 2022 Jan 28;17(1):e0261940.

Abstract

Objectives

To assess the cost-effectiveness of various combinations of urate lowering therapy (ULT) and anti-inflammatory treatment in the management of newly diagnosed gout patients, from the Dutch societal perspective.

Methods

A probabilistic patient-level simulation estimating costs and quality-adjusted life years (QALYs) comparing gout and hyperuricemia treatment strategies was performed. ULT options febuxostat, allopurinol and no ULT were considered. Flare treatments naproxen, colchicine, prednisone, and anakinra were considered. A Markov Model was constructed to simulate gout disease. Health states were no flare, and severe pain, mild pain, moderate pain, or no pain in the presence of a flare. Model input was derived from patient level clinical trial data, meta-analyses or from previously published health-economic evaluations. The results of probabilistic sensitivity analyses were presented using incremental cost-effectiveness ratios (ICERs) and summarized using cost-effectiveness acceptability curves (CEACs). Scenario analyses were performed.

Results

The ICER for allopurinol versus no ULT was €1,381, when combined with naproxen. Febuxostat yielded the highest utility, but also the highest costs (€4,385 vs. €4,063 for allopurinol), resulting in an ICER of €25,173 when compared to allopurinol. No ULT was not cost-effective, yielding the lowest utility. For the gout flare medications, comparable effects on utility were achieved. Combined with febuxostat, naproxen was the cheapest option (€4,404), and anakinra the most expensive (€4,651). The ICER of anakinra compared to naproxen was €818,504. Colchicine and prednisone were dominated by naproxen.

Conclusion

Allopurinol and febuxostat were both cost-effective compared to No ULT. Febuxostat was cost-effective in comparison with allopurinol at higher willingness-to-pay thresholds. For treating gout flares, colchicine, naproxen and prednisone offered comparable health economic implications, although naproxen was the favoured option.

Introduction

Gout is an inflammatory response to the presence of hyperuricemia induced monosodium urate (MSU) crystals within the synovial fluid of joints and tissues. It is the most common cause of inflammatory arthritis in men, and reports have shown the burden of gout to be rising [101]. Gout attacks are characterized by rapid onset of severe pain and may have a considerable impact on patient's ability to work and function in other social roles [102,103]. Typically, gout attacks resolve within 5-7 days with effective anti-inflammatory treatment. Recurrent attacks, and the development of chronic, inflammatory gout, may be prevented by effective urate lowering therapy (ULT) aimed at lowering serum urate (SUA) levels below the saturation point for crystal formation [104]. In light of the increasing burden of gout, the importance of optimizing treatment and management of gout at various levels, including patient, community and national, is emphasized.

Various safe and effective anti-inflammatory therapies are available for the treatment of both gout attacks and hyperuricemia. Allopurinol and febuxostat are currently recommended first-line ULT agents [105]. Colchicine, non-steroidal anti-inflammatory drugs (NSAIDs) and glucocorticosteroids are all first-line treatment options for treating gout flares [106]. Besides these traditional synthetic medications, targeted biological medications, in particular interleukin-1 (IL-1) inhibitors, including anakinra and canakinumab, have been investigated in recent clinical trials for treating gout flares [24,107]. IL-1 inhibition is currently recommended as a second-line treatment option for managing gout flares [105,106].

Due to the high and increasing prevalence of gout, and the introduction of novel treatment options such as the relatively expensive IL-1 inhibitors, health economic implications are important to consider when deciding on optimal treatment approaches for patients with this disease [108]. Health economic decision models that have thus far been developed to support such decision making are mainly concerned with the comparison of various ULTs [109–111]. Although some models do account for gout flares by assigning disutility, the effects and costs of anti-inflammatory treatments are not explicitly considered in addition to or instead of ULT. However, with the introduction of new, more costly and potentially more effective treatments for treating gout flares, simultaneously evaluating outcomes of ULT and anti-inflammatory medications becomes more relevant. In the present study we introduce a new modelling framework for gout, in which the costs and effects of treatment strategies with continuous ULT and anti-inflammatory medications for gout flare can be assessed in newly diagnosed patients. This model compares the costs and effects of various commonly recommended and administered ULT and anti-inflammatory treatments from a Dutch societal perspective.

Methods

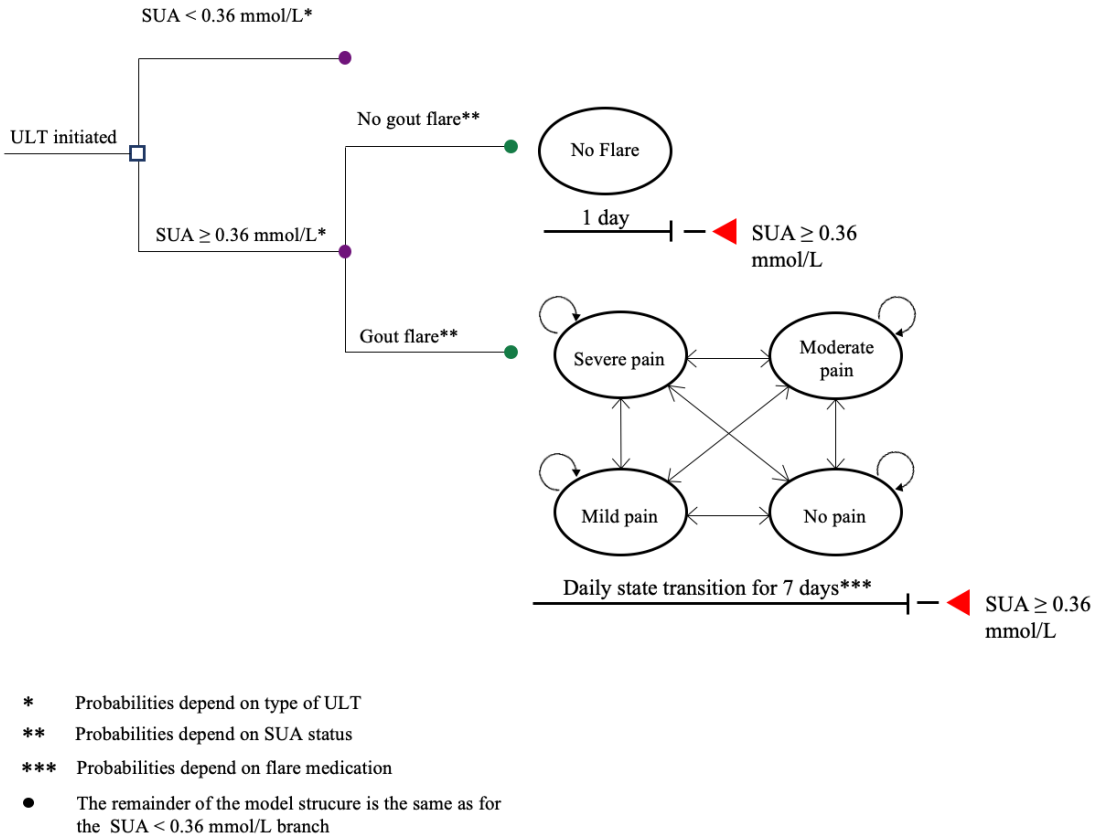
For this study, model-based cost-effectiveness analysis was performed from the societal perspective, in which first-line ULT agents for hyperuricemia (i.e. allopurinol, febuxostat, no ULT), as well as first-line (i.e. colchicine, naproxen and prednisone) and second-line (i.e. anakinra) treatment options for gout flares were compared. Medication costs, other healthcare costs and

productivity loss were included. The Dutch Willingness To Pay (WTP) threshold is not strictly defined and can be calculated based on disease burden. The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement was followed in reporting the results of this cost-effectiveness analysis (See **Supplementary Material 1**).

Markov Model

A Markov Model (TreeAge™) was developed to simulate and compare outcomes of various ULT and anti-inflammatory treatment combinations for hyperuricemia and gout flares, respectively (Figure 1). For the present study the time horizon was one year with a cycle length of one day. Due to the nature of the available data for anakinra and the absence of long-term effects of gout in the model a much longer horizon would not be appropriate. Running a longer-term model with the lack of long-term data would lead to serious omissions. The model considered the one-year course of newly diagnosed gout patients, receiving treatment for their gout flare, and who initiated ULT while experiencing a gout flare, reflecting a care path commonly applied for these gout patients in clinical practice. Upon entry in the model, patients were assigned a fixed dose of ULT, with either allopurinol (at 200 mg or 300 mg), or febuxostat 80 mg, or no ULT, based upon available data from the literature. The probability that the patient achieved the SUA target, defined as achieving a SUA level < 0.36 mmol/L, depended on the specific ULT. Patients were not able to switch after this first division had been made, thus, they could not go from SUA < 0.36 mmol/L to any other branch for the one-year duration of this model. After it had been determined if a simulated patient would have SUA levels on target (< 0.36 mmol/L) or not on target (≥ 0.36 mmol/L) for the duration of the simulation, patients experienced a daily risk of having a gout flare. Patients with SUA levels not on target (in the SUA ≥ 0.36 mmol/L branch) had a higher daily flare risk than patients in the on-target branch (SUA < 0.36 mmol/L). Patients in both branches were able to experience flares. Patients that did not reach the target did experience more and more frequent flares in this model. When a gout flare was triggered, patients transitioned for seven days between four mutually exclusive pain states (i.e. no pain, mild pain, moderate pain, severe pain) according to transition probabilities defined for each gout flare treatment option. Treatment options included in the model for gout flares were colchicine, naproxen, prednisone and anakinra. Dosages were in line with the dosages as used in the clinical studies used as data sources [24,112]. No switching between ULT drugs, or medication for gout flares, were allowed in the model during the time horizon. Costs and quality-adjusted life years (QALYs) over one year were recorded for all strategies.

Figure 1: Model Structure



ULT, urate lowering therapy; SUA, serum urate. * Probabilities depend on type of ULT; ** Probabilities depend on SUA status; *** Probabilities depend on gout flare medication; • The remainder of the model structure is the same as for the *SUA < 0.36 mmol/L*-branch

Analyses

We employed probability sensitivity analyses (PSA) with 2000 x 200 model replications to take uncertainty around the point estimates of the model parameters into account. The results were summarized using cost-effectiveness acceptability curves (CEACs). Costs were discounted at 4%, utility was discounted at 1.5%, in concordance with the Dutch costing manual [71]. The WTP thresholds in the Netherlands are not explicitly set but lie between €10,000 and €80,000 per QALY [113].

Model Inputs

All parameters used as input for the model, as well as the data source from which they were estimated, are listed in Table 1. Transition matrices were used in the model between pain states, these can be found in the supplemental material 2: table 1. Uncertainty regarding elements of the various transition matrices were expressed using Dirichlet distributions. For other input parameters, various distributions were fitted to the observed data. Chi-squared and Anderson Darling fit statistics were used to evaluate goodness of fit.

Table 1: Model inputs.

Parameter	Point estimate	Probability distribution	Source
Probability SUA < 0.36 mmol/L			
Allopurinol 200 mg	0.457	Beta (μ: 0.455, σ: 0.045)	[114]
Allopurinol 300 mg	0.480	Beta (μ: 0.478, σ: 0.027)	[115–117]
Febuxostat 80 mg	0.729	Beta (μ: 0.730, σ: 0.013)	[115,117–120]
No ULT	0.000	n/a	[121]
Daily flare probability			
SUA on target with ULT	0.000716	Beta (μ: 0.999, σ: 0.002)	[111]
SUA not on target with ULT	0.001222	Beta (μ: 0.998, σ: 0.002)	[111]
SUA not on target with no ULT	0.001637	Beta (μ: 0.998, σ: 0.002)	[111]
Quality adjusted life days			
No pain	0.86	Beta (α: 16.325, β: 3.076)	[24]
Mild pain	0.77	Beta (α: 10.942, β: 3.177)	[24]
Moderate pain	0.70	Beta (α: 12.696, β: 5.329)	[24]
Severe pain	0.61	Beta (α: 19.817, β: 12.877)	[24]
Daily other costs (in €s)			
No pain	19.95	Exp/g (λ: 1.749 /α: 0.317, β: 3.623)[1]	[24,71]
Mild pain	32.39	Exp/g (λ: 1.081 /α: 0.493, β: 2.135)	[24,71]
Moderate pain	58.08	Exp/g (λ: 0.599 /α: 0.434, β: 4.149)	[24,71]
Severe pain	134.32	Exp/g (λ: 0.259 /α: 0.291, β: 10.769)	[24,71]
Daily drug costs (in €s)			
Colchicine	0.61	n/a	[122]
Naproxen	0.21	n/a	[122]
Prednisone	0.26	n/a	[122]
Anakinra	33.4	n/a	[122]
Allopurinol	0.13	n/a	[122]
Febuxostat	1.03	n/a	[122]

ULT: Urate lowering therapy, SUA: serum urate

[1] λ-parameter refers to WPAI (exponential distribution) and α and β refer to the ZoCo (gamma distribution)

Efficacy of ULT

ULT success was defined as achieving SUA level < 0.36 mmol/L. This SUA target level is recommended by guidelines, supported by reports that have shown that SUA levels below the target level of 0.36 mmol/L are associated with a decreased risk for gout flares [106,123]. To generate model input, a meta-analysis was performed of ULT clinical trials, in which achieving the SUA target of < 0.36 mmol/L was one of the endpoints. The indirect adjusted comparison method, using febuxostat 40 mg as the reference treatment was used to obtain efficacy estimates and associated standard errors adjusted for study specific factors [124]. For all the placebo arms in the meta-analysis, the percentage of patients achieving the target was zero percent. Therefore, for all treatment strategies in which patients do not use ULT it was assumed that no patient achieved the SUA target.

Flare probabilities

Daily flare probabilities were calculated from data derived from a previous health economic model by Jutkowitz et al. 2014 [111]. In that paper, annual flare probabilities were given for patients on ULT with controlled SUA (< 0.36 mmol/L), for patients on ULT with uncontrolled SUA (≥ 0.36 mmol/L), as well as annual flare probabilities for patients not on ULT, with uncontrolled SUA (≥ 0.36 mmol/L).

Efficacy of flare treatment

Health states for patients experiencing a gout flare were defined using four pain states (i.e. no pain, mild pain, moderate pain, severe pain), derived from a 4-point pain rating scale that is commonly used as a primary endpoint in gout clinical trials. Inverse variance weighted pain transition probabilities and their standard errors for naproxen and prednisone were obtained by pooling seven day [24]. And 90 hours [112] follow up data from two clinical trials. For colchicine and anakinra, the probabilities were obtained from the seven days follow up data of a single trial [24]. To avoid empty cells in the transition matrices, due to data sparsity, a Bayesian approach was used in which a transition matrix with 0.5 for each cell (i.e. noninformative prior) was combined with the observed transition frequencies [125]. Dirichlet distributions were then fitted on the resulting posterior distribution of transition probabilities in the PSA.

Utilities

Utility weights were estimated for each of the four pain related health states using data obtained from the study by Janssen et al. 2019 [24]. The values of utility for each pain state were calculated from the SF-6D. Since a Dutch tariff is unavailable, the SF-6D health stages were valued with the UK tariff.

Costs

We included costs related to gout drug use, healthcare resource utilization and work productivity loss due to gout and other reasons, in euros using Dutch price indices. The costs per day for each pain state were determined for all cost variables. The medication costs for the appropriate dosages

were derived from the official Dutch list prices [122], and the cost per hour of foregone labour was retrieved from a report by The Netherlands Organizaton for Scientific Research [72].

Healthcare resource utilization for each pain state was estimated using patient-reported questionnaires obtained during a gout clinical trial, wherein patients were asked to report the number of visits to general practitioners, outpatient clinics for specialized caregivers, paramedical caregivers, but also the amount of household care that was used, and diagnostic tests (i.e. echo, CT scan) undergone [24]. The costs for each of these items were obtained from the 2015 Dutch Costing manual and corrected to the appropriate costing year [71]. Outliers, defined as costs deviating more than three standard deviations from the mean, were removed from the healthcare utilization data.

Scenario analyses

Scenario analyses were performed for gout patients who did not experience a gout flare at model entry, and for gout patients with severe gout. For the latter, the daily flare chances were successively increased.

Results

As the model used in this study incorporated treatment options for gout flares and hyperuricemia, there were two types of medication (i.e. anti-inflammatory treatment and ULT) compared in this study, resulting in many possible combinations. This section will first focus on different ULT medications combined with naproxen. Hereafter, the cost-effectiveness PSA results of different gout flare medication combined with different ULT medication types, will be presented.

PSA Results ULT medication with naproxen as anti-inflammatory treatment

Table 2 shows the PSA results of three ULT medication types combined with naproxen. No ULT was the cheapest option at €4,031.19. Allopurinol yielded more utility at a slightly higher cost of €4,063.94. Compared with no ULT, allopurinol was cost-effective with an incremental cost-effectiveness ratio (ICER) of €1,381.27. The most expensive option was febuxostat (€4,385.40), also yielding the highest utility. Compared to allopurinol, febuxostat had an ICER of €25,173. At a WTP-threshold of €25,173, this would be considered cost-effective in The Netherlands.

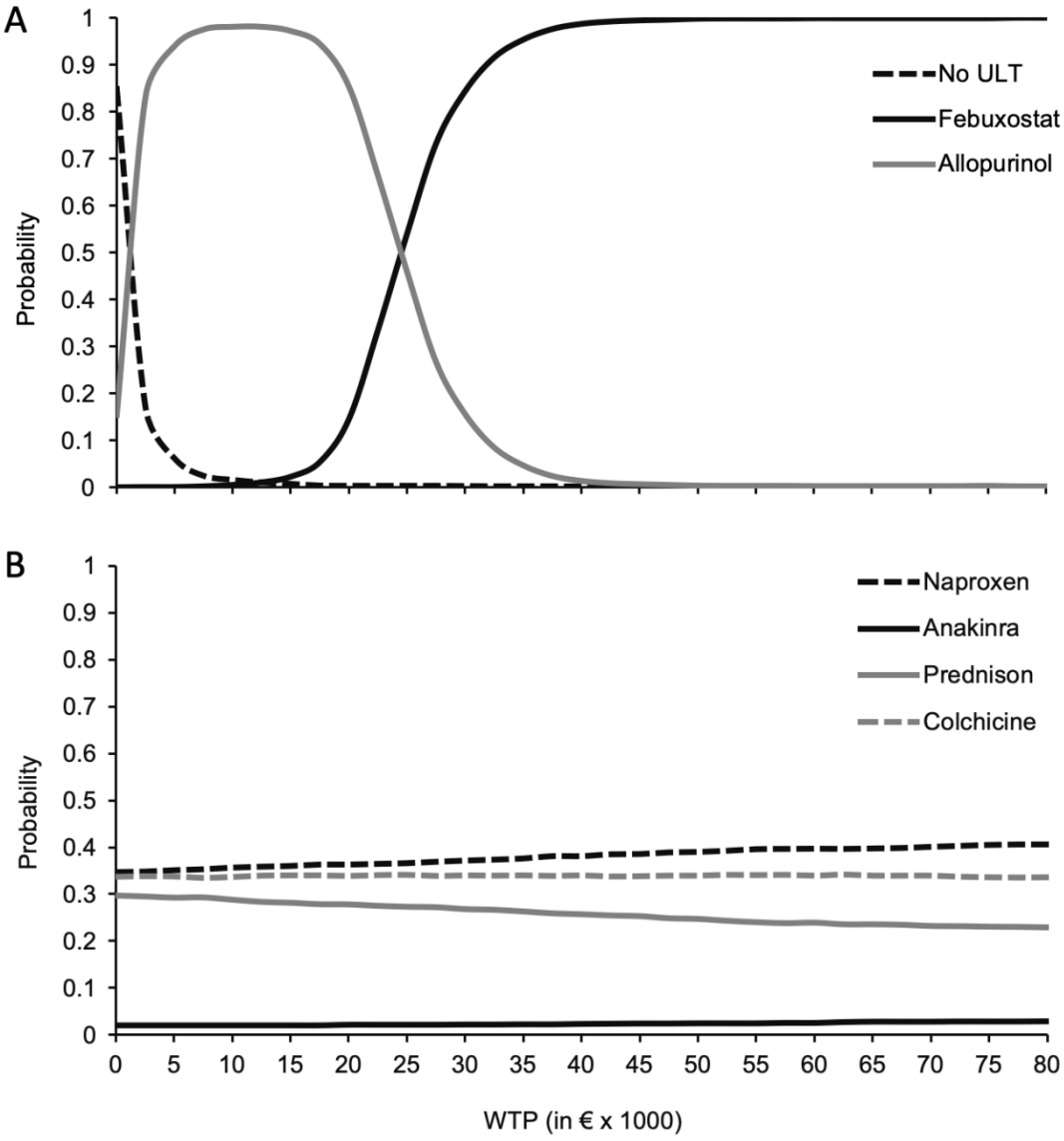
Table 2: Probabilistic Sensitivity Analysis for urate lowering therapy combined with naproxen.

	Costs (€s)	Effects (QALY)	ΔC (€s)	ΔE (QALY)	ICER (Δ€/Δ QALY)
No ULT + Naproxen	4,031.19	0.78877	-	-	-
Allopurinol + Naproxen	4,063.94	0.81248	32.75	0.02371	1,381.27*
Febuxostat + Naproxen	4,385.40	0.82525	321.46	0.01277	25,173.06**

QALY = Quality adjusted life years; ICER = incremental cost-effectiveness ratio; ULT = urate lowering therapy; *ICER for allopurinol vs NO ULT; ** ICER for febuxostat vs allopurinol

Figure 2 displays the cost-effectiveness acceptability curves (CEACs). Overall, the differences in utility between acute flare medications (anakinra, colchicine, naproxen, and prednison) are very small. This is also reflected in the CEAC. The 0 point at the y-axis reflects the probability of each of the strategies being cost-saving, whereafter all points on the graph reflect the certainty with which we can say that the ICER is acceptable. The first panel (A) of Figure 2 shows that at WTP-thresholds below €25,173, allopurinol is the preferable option. At a WTP-threshold of approximately €25,000, there is a ~50% chance that allopurinol is cost-effective. At WTP-thresholds above €25,173, the probability with which febuxostat is cost-effective grows to ~100%, at a WTP of approximately €35,000. Results and implications were similar when combining the ULT medication with any of the other flare medications. The second panel of figure 2 (B) shows the CEAC of four strategies, febuxostat combined with the four anti-inflammatory medications. Over the full range of WTP threshold, febuxostat + naproxen has the highest probability of being cost-effective, ranging from 36% to 39%. Febuxostat combined with colchicine remains around 35% and febuxostat + prednisone ranges from 30% to 24%.

Figure 2: Cost-effectiveness acceptability curves (CEAC)



ULT. urate lowering therapy; WTP, willingness-to-pay. Panel A displays the CEAC for different ULT combined with naproxen as the anti-inflammatory agent. Panel B displays the CEAC for different anti-inflammatory treatment options combined with febuxostat as the ULT.

PSA Results Comparing anti-inflammatory treatments with febuxostat as ULT

The different anti-inflammatory treatments had comparable effects on utility over 1 year when they were combined with febuxostat as ULT. PSA results showed that when naproxen was combined with febuxostat, patients accrued an estimated 0.81 QALYs over the course of 1 year. Larger differences between flare medication were seen in costs. Naproxen combined with febuxostat was the cheapest option at an estimated cost of €4,404,- per year. Colchicine and

prednisone were slightly more expensive at lower accrued utility. Anakinra was the most expensive anti-inflammatory medication. Anakinra did yield higher utility than naproxen, but at an ICER of €818,504 comparing anakinra to naproxen, both combined with febuxostat would not be considered cost-effective in the Netherlands with WTP-thresholds ranging from €10,000 to €80,000.

The small differences in QALYs gained between flare medications (anakinra, colchicine, naproxen, and prednisone) and higher incremental costs of anakinra over the other medications were also reflected in the cost-effectiveness acceptability curves (CEAC) (Figure 2). Naproxen had the highest probability of being cost-effective across the full range of WTP-thresholds, ranging around 33%. As the WTP-threshold increases to €80,000, the probability of naproxen being cost-effective increased to ~42%. However, both colchicine and prednisone had only slightly lower probabilities of being cost-effective than naproxen. Colchicine is stable across the WTP-threshold range at 33% chance of being cost-effective. Prednisone had a lower probability of being cost-effective and it decreased as the WTP-threshold increases. **Table 3** also shows results when flare medication is combined with allopurinol and No ULT. In both cases, naproxen is the favourable option.

Table 3: Probabilistic Sensitivity Analysis for anti-inflammatory treatment

	Costs (€)	Effect (QALY)	ΔC	ΔE	ICER
Allopurinol					
Naproxen	4,051.37	0.812493	-	-	-
Colchicine	4,066.48	0.812356	15.11	-0.00014	Dominated
Prednisone	4,074.27	0.812356	22.9	-0.00014	Dominated
Anakinra	4,299.31	0.81274	247.94	0.00025	1,003,805.67
Febuxostat					
Naproxen	4,404.72	0.82288	-	-	-
Colchicine	4,426.91	0.82274	22.19	-0.00014	Dominated
Prednisone	4,424.72	0.82277	20.00	-0.00011	Dominated
Anakinra	4,651.09	0.82318	246.37	0.00030	818,504.98
No ULT					
Naproxen	4,012.99	0.788082			-
Colchicine	4,039.99	0.787918	27.00	-0.00016	Dominated
Prednisone	4,040.82	0.757973	27.83	-0.03011	Dominated
Anakinra	4,297.02	0.788356	284.03	0.00027	1,036,605.84

QALY = Quality adjusted life years; ICER = incremental cost-effectiveness ratio; ULT = urate lowering therapy; *ICER for allopurinol vs NO ULT; ** ICER for febuxostat vs allopurinol

Scenario analyses

In the scenario of patients without flare upon entry, flares per year were lower compared to the base case analysis. This resulted in slightly higher utility and lower costs, but the conclusion remained the same. The second scenario concerns patients with a higher daily probability of starting a flare. Overall, this resulted in lower utility and higher costs. Again, the implications did not change. Naproxen remained the most favourable option. The preferred ULT option still depends on the WTP-threshold that would be set in gout, but it remained clear that no ULT was not a preferable option (See Supplemental material 3).

Discussion

A health economic model was developed for evaluating the costs and effects related to gout treatment strategies that simultaneously covers anti-inflammatory agents for gout flares (i.e. colchicine, naproxen, prednisone, anakinra) and ULT options for hyperuricemia (i.e. allopurinol, febuxostat, or no ULT).

The results of our comparison of ULT strategies suggest that strategies in which no ULT is used would not be considered cost-effective at any WTP threshold that is customary in the Netherlands. This finding supports, from a health-economic point of view, the 2016 updated EULAR guidelines,

which was the first to emphasize that ULT should be considered and discussed with every patient from the first presentation of gout with a definite diagnosis [106]. It should be noted that due to the relatively small incremental QALYs between strategies, costs play a rather significant role in the outcome of these analyses. Our results further show that which specific ULT yields the highest net benefit depends on the WTP threshold. In the Netherlands, the WTP threshold ranges from €10,000 to €80,000 and depends on the 'burden of disease', estimated using the proportional shortfall method. The ICER of febuxostat compared to allopurinol is €25,173.06 and is thus quite close to the WTP-threshold set for the lowest disease burden category, which is up to €20,000 euro per QALY [113]. Although not yet explicitly defined, the disease burden of the population of gout patients considered in this study could be expected to fall in the lowest category defined by the National Healthcare Institute [113]) due to its episodic pattern with longer periods of no attacks. However, there are various methods of calculating burden of disease [126]. With respect to the cost-effectiveness of different ULT, this would indicate that using allopurinol is the preferable option. However, the disease burden of gout varies substantially with severity. For example, health-related quality of life of patients with difficult to treat, chronic gout was found to be similar to that of patients with active rheumatoid arthritis [127]. Since the proportional shortfall weighted burden of rheumatoid arthritis corresponds to the highest disease burden category [128], a WTP threshold of €80,000 might be applied to the population of patients with severe gout. This would suggest that febuxostat may be preferable to allopurinol in the treatment of chronic gout. However, as the results of our study were not based on patients with severe gout, this would need to be investigated further in future studies.

The cost-effectiveness of various ULT monotherapies has been compared in several previous, model-based studies [109–111]. In all cases, these studies considered ULT only and, either did not consider the impact of flares on quality of life or used a disutility to account for flares. The current paper specifically focusses on combinations of ULT and flare medication. Furthermore, utility and costs weights were attached to various SUA level related health states. By contrast, in our model utility and costs are mainly determined by the current level of pain experienced by the patient, with a disutility for patients not reaching the SUA target. Utility was mainly determined by flare duration and intensity. This choice was motivated by the consistent findings in previous studies that pain is strongly related to health-related quality of life of gout patients, whereas mixed findings were reported with respect to the relationship between SUA levels and quality of life [129]. The assumption that lowering SUA levels would produce utility gains independently of gout flares was also considered implausible in a recent NICE single technology appraisal of cost-effectiveness evidence in favour of febuxostat, since gout is usually asymptomatic in between flares [130]. Nevertheless, it is interesting to note that despite these differences in model structure, all three studies found results roughly consistent with ours in that applying any ULT was found to be cost-effective relative to no ULT. Additionally, model input was limited by data availability. For example, higher dose allopurinol treatments are rarely administered in daily clinical practice due to intolerance and has not provided an adequate amount of data. It can be hoped for that with the upswing of registries and technological advancements more knowledge and daily clinical practice data becomes available. Furthermore, febuxostat was consistently associated with both higher costs and higher effectiveness compared with allopurinol in the previous studies.

Several more studies evaluated cost-effectiveness of various ULT sequences [131–133]. However, our analyses solely focused on monotherapies since this has been shown to be the most common treatment pattern in clinical practice [111,134]. It should be noted though that various international guidelines currently recommend titrating allopurinol dosages up to 900 mg/day. It seems likely that a higher percentage of patients would be able to reach the SUA target at higher allopurinol dosages. Unfortunately, no suitable data was identified in our literature review to be able to assess this treatment strategy in our model.

To the best of our knowledge, the current study is also the first model-based study to examine cost-effectiveness of anti-inflammatory treatments for gout attacks. Results of our study reveal that naproxen was the favourable treatment at any WTP-threshold, in combination with any of the ULT, although overall differences in cost-effectiveness between conventional treatment strategies remained small. In addition, our results showed that treatment with anakinra, although accruing slightly higher health outcomes after one year compared to conventional treatments, was not cost-effective, primarily driven by its high costs per treatment. Costs over one year for strategies including anakinra were approximately €200,- higher than the other gout flare medications. Although this is a smaller difference than what the difference in absolute drug prices between anakinra and, for example, naproxen would suggest, our findings do not support a role for anakinra as a first line treatment in the overall gout population.

The current study had some limitations. First, the amount of data used to estimate the pain transition probabilities for each gout flare treatment option, was limited. This resulted from the need to have access to patient level data to populate the model. In particular, data with regard to anakinra and colchicine were based on a single randomized controlled trial. The resulting uncertainty about the relative effectiveness of the different treatments may have undermined our ability to differentiate the efficacy of different anti-inflammatory treatments. Second, the occurrence of (serious) adverse events and their associated costs and consequence on utilities, were not included in the model. This also applies for using prophylaxis when initiation ULT as recommended by gout guidelines. Next, insufficient data was available to consider running a longer time horizon for this model. Information on efficacy and safety of gout treatments in the longer term was available, however, for using that data in modelling studies like the current paper it would have to have been linked to quality of life or utility data, which was not an option with the current existing data. The one-year time horizon for this model allowed to focus on the effects of gout in a newly diagnosed population. For patients that experience relatively many flares, quality of life mostly depends on their health states (and utility) during those flares. While a longer term model would certainly also be interesting and necessary for decision makers, long-term effects are not within the scope of this model. A longer time horizon would mean that patients are more stable on their ULT, experience less flares and may choose to discontinue or stop ULT and rarely need flare medication. The scope of the current research has specific attention for utilities and costs during flares, therefore the time horizon has been limited. A longer horizon would demand a different focal point. When looking at a longer term model, several more events could be included. In an ideal situation with a richness of data sources and unlimited modelling options, it would be fascinating to be able to include SAE's, ULT sequences, medication discontinuation, and other long-term gout events. Insufficient data was available for us to consider the IL-1 inhibitor,

canakinumab, for the treatment of gout flares, or second-line ULT agents as pegloticase and lesinurad, which has just recently been approved by the Food and Drug Administration in combination with allopurinol [135]. However, none of these drugs are likely to become first line treatment options for gout and hyperuricemia in the near future.

Conclusions

In conclusion, the findings of our study show that ULT, with either allopurinol or febuxostat, are cost-effective first-line ULT agents for treating hyperuricemia. For the treatment of gout flares, conventional first-line treatments (i.e. colchicine, naproxen, prednisone) had similar health economic implications, of which naproxen had the most favourable costs and effects profile.

Supplemental Material

Table 1: Transition matrix for naproxen

	End pain state (%)			
Start pain state	No	Mild	Moderate	severe
No	91.46	6.09	1.21	1.21
Mild	13.28	80.44	4.42	1.85
Moderate	0.87	45.61	43.86	9.65
Severe	2.17	14.13	36.96	46.74

Table 2: No flare upon entry full results

	Costs (€)	Effects (QALY)	ΔC (€)	ΔE (QALY)	ICER (ΔC/ΔE)
Febuxostat					
Naproxen	4,146	0.82484	-	-	-
Colchicine	4,152	0.82481	5.95	-0.00003	Dominated
Prednisone	4,147	0.82481	0.42	-0.00003	Dominated
Anakinra	4,217	0.82482	71.19	-0.00002	Dominated
Allopurinol					
Naproxen	3,561	0.82436	-	-	-
Colchicine	3,566	0.82426	4.77	-0.00010	Dominated
Prednisone	3,563	0.82431	1.54	-0.00005	Dominated
Anakinra	3,643	0.82431	81.70	-0.00005	Dominated
No ULT					
Naproxen	3,885	0.78639	-	-	-
Colchicine	3,894	0.78632	8.49	-0.00007	Dominated
Prednisone	3,893	0.78634	7.67	-0.00005	Dominated
Anakinra	3,989	0.78636	103.91	-0.00003	Dominated
ULT comparison					
No ULT + Naproxen	3,885	0.78639			
Allopurinol + Naproxen	3,561	0.82436	-324	0.03797	Dominating*
Febuxostat + Naproxen	4,146	0.82484	585	0.00048	1,218,750**

QALY = Quality adjusted life years; ICER = incremental cost-effectiveness ratio; ULT = urate lowering therapy; *ICER for allopurinol vs No ULT; ** ICER for febuxostat vs allopurinol

Table 3: Increase flare chance full results

	Costs (€)	Effects (QALY)	ΔC (€)	ΔE (QALY)	ICER (ΔC/ΔE)
Febuxostat					
Naproxen	4,389	0.82507	-	-	-
Colchicine	4,417	0.82492	28.25	-0.00015	Dominated
Prednisone	4,406	0.82493	17.05	-0.00014	Dominated
Anakinra	4,642	0.82535	252.85	0.00027	927,118
Allopurinol					
Naproxen	4,069	0.80843	-	-	-
Colchicine	4,099	0.80829	29.22	-0.00013	Dominated
Prednisone	4,090	0.80828	20.23	-0.00014	Dominated
Anakinra	4,336	0.80872	266.45	0.00030	896,226
No ULT					
Naproxen	4,043	0.78905	-	-	-
Colchicine	4,075	0.78889	32.34	-0.00015	Dominated
Prednisone	4,076	0.78890	33.02	-0.00015	Dominated
Anakinra	4,331	0.78934	288.38	0.00029	978,458
ULT Comparison					
No ULT + naproxen	4,043	0.78905	-	-	-
Allopurinol + naproxen	4,069	0.80843	26	0.01938	1,342*
Febuxostat + naproxen	4,389	0.82507	320	0.01664	19,231**

QALY = Quality adjusted life years; ICER = incremental cost-effectiveness ratio; ULT = urate lowering therapy; *ICER for allopurinol vs No ULT; ** ICER for febuxostat vs allopurinol

Section II

Section II – Responsiveness of Instruments in Measuring Deterioration in Health Status


Chapter

5

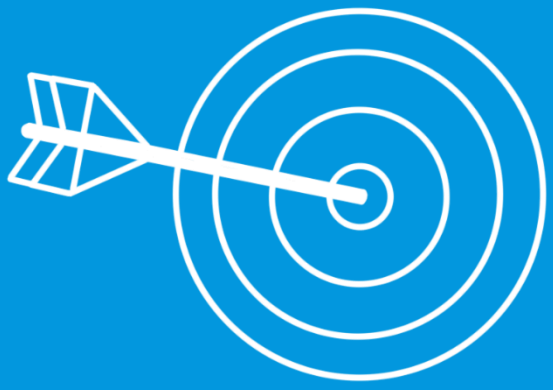

Chapter 5 – Detecting Disease Worsening in Rheumatoid Arthritis Patients

Submitted for publication as: Limited responsiveness of both generic and disease-specific patient reported outcome measures in detecting disease worsening during deterioration in Rheumatoid Arthritis Patients: need for improvements

Van de Laar, C.J., ten Klooster, P.M., Oude Voshaar, M.A., Van de Laar, M.A.F.J., Uyl-De Groot, C.A.
Submitted for publication

Embargo

Section



III

Section III – Evaluating Real-World Outcomes of Treatments in Rheumatology



Chapter 6



Chapter 6 – A Pragmatic Trial Comparing Baricitinib versus TNF Inhibitors

Published as: PERFECTRA: a pragmatic, multicentre, real-life study comparing treat-to-target strategies with baricitinib versus TNF inhibitors in patients with active rheumatoid arthritis after failure on csDMARDs

Van de Laar CJ, Voshaar MA, Ten Klooster P, Tedjo DI, Bos R, Jansen T, Willemze A, Versteeg GA, Goekoop-Ruiterman YP, Kroot EJ, van de Laar M. PERFECTRA: a pragmatic, multicentre, real-life study comparing treat-to-target strategies with baricitinib versus TNF inhibitors in patients with active rheumatoid arthritis after failure on csDMARDs. RMD open. 2024 May 1;10(2):e004291.

Abstract

Objective

To compare the effectiveness of a strategy administering baricitinib versus one using TNF-inhibitors (TNFi) in patients with rheumatoid arthritis (RA) after csDMARDs failure in a real-life treat-to-target (T2T) setting.

Methods

Biological and targeted synthetic DMARD (b/tsDMARD) naïve RA patients with disease duration ≤5 years without contraindications to b/tsDMARD were randomized to either TNFi or baricitinib when csDMARD failed to achieve disease control in a T2T setting. Changes in clinical and patient-reported outcome measures (PROMs) were assessed at 12-week intervals for 48 weeks. The primary endpoint was non-inferiority, with testing for superiority if non-inferiority is demonstrated, of baricitinib strategy in the number of patients achieving ACR50 response at 12 weeks. Secondary endpoints included DAS28-CRP <2.6, changes in PROMs and radiographic progression.

Results

A total of 199 patients (TNFi, n=102; baricitinib, n=97) were studied. Both study groups were similar. Baricitinib was both non-inferior and superior in achieving ACR50 response at week 12 (42% vs. 20%). Moreover, 75% of baricitinib patients achieved DAS28-CRP <2.6 at week 12 compared to 46% of TNFi patients. On secondary outcomes throughout the duration of the study, the baricitinib strategy demonstrated comparable or better outcomes than TNFi strategy. Although not powered for safety, no unexpected safety signals were seen in this relatively small group of patients.

Conclusion

Up to present, in a T2T setting, RA patients failing csDMARDs have two main strategies to consider, JAKi versus bDMARDs (in clinical practice, predominantly TNFi). PERFECTRA suggested that starting with baricitinib was superior over TNFi in achieving response at 12 weeks and resulted in improved outcomes across all studied clinical measures and PROMs throughout the study duration in these patients.

Introduction

In recent decades, outcomes for rheumatoid arthritis (RA) have improved significantly, largely due to biological therapies and the treat-to-target (T2T) paradigm [179,180]. Despite these advances, a considerable number of patients in real-world clinical practice fail to achieve sustained disease control. Targeted synthetic disease-modifying antirheumatic drugs (tsDMARDs) have emerged as a new disease modifying treatment option [181–184]. However, their effectiveness in comparison with other options within the current T2T strategies in the real world can still be better understood. Additionally, there is a need for real-world evidence of these treatment options within the T2T framework.

Baricitinib is an oral Janus Kinases inhibitor (JAKi) selectively for JAK-1 and -2 available for treatment of RA [181,185,186]. Even though EULAR guidelines present tsDMARDs, including JAKi such as baricitinib, as an treatment option next to biological disease-modifying antirheumatic drugs (bDMARD) after methotrexate (MTX) failure to reach the target with respect to disease activity: remission and low disease activity, [180]. Physicians' extensive clinical experience with bDMARDs and specifically tumor necrosis factor inhibitors (TNFi), causes these to be more commonly used in practice. Other factors like the availability of several TNFI biosimilars and costs, obviously also play a role in usage in practice.

However, tsDMARDs offer several clinically relevant benefits for patients compared to bDMARDs (TNFi), including convenient mode of administration, short half-life, and improved suitability for monotherapy. RCTs are extremely valuable in assessing efficacy and effectiveness of the compound of interest. However, due to strict in- and exclusion criteria and other factors they may not always fully reflect how care is being conducted in all situations that the real world may bring. To add to the evidence that formal RCTs generate and to create a fuller picture of all scenarios, real world studies are incredibly important and can complement results from RCTs to generate more information and further improve patient care. It is therefore important to study treatment options and strategies in relevant real-life settings, like in settings where T2T is fully implemented [187,188]. Obviously, compliance with safety precautions and recommendations surrounding treatment of immune-mediated inflammatory diseases with JAKi have to be fully taken into account as outlined in Nash et al.[189].

PERFECTRA was designed as a pragmatic trial to inform clinical practice on the effectiveness of a strategy starting with the JAKi baricitinib compared with a strategy starting with a TNFi after csDMARDs failure in a real-life T2T setting. The pragmatic design of PERFECTRA yields several benefits, it better incorporates aspects of the population and its characteristics for which an intervention is intended and thus evaluates effectiveness in real-life situations whereas results from formal RCTs can be limited in generalizability [190]. The daily clinical practice setting and limited in- and exclusion criteria maximize applicability of the results of PERFECTRA. In this study, we report the findings of the 48-week multicentre randomized, open-label, pragmatic real-world PERFECTRA trial. The primary goal of PERFECTRA was to establish the non-inferiority of the tsDMARD baricitinib to TNFi in terms of American College of Rheumatology 50 (ACR50) response at 12 weeks. If non-inferiority was confirmed, assessment of superiority of baricitinib strategy at

12 weeks was included. Secondary objectives encompassed the comparison of patient-reported outcomes (PROMs), safety assessments, and radiological damage over the course of 48 weeks.

Methods

Ethical statement

The patients included in PERFECTRA gave written informed consent prior participation. The study protocol has been approved by the Medical Ethics Committee East Netherlands and is registered in the International Clinical Trials Registry Platform (https://trialsearch.who.int; NL7547).

Study Design

The investigator-initiated PERFECTRA study was a 48-week multi-centre randomized, open label, pragmatic, real-world noninferiority (including superiority) trial designed for 200 patients with active RA, despite adequate dosage of csDMARD. Included patients were treated open label, according the T2T principle [180] to either a strategy starting with TNFi (any TNFi as indicated and reimbursed for RA treatment in the Netherlands) or a treatment strategy starting with baricitinib.

Patients

Inclusion criteria were a clinical diagnosis of RA, active disease at the discretion of the rheumatologist, former treatment according to T2T principles (i.e. past treatment decisions informed by disease activity measurements), and previous use of at least one conventional synthetic disease-modifying antirheumatic drug (csDMARD). Exclusion criteria included disease duration longer than 5 years, previous treatment with any bDMARD or tsDMARD, contraindications for TNFi and baricitinib, failure to provide written informed consent, or a refusal to use effective contraceptive during the study period when applicable.

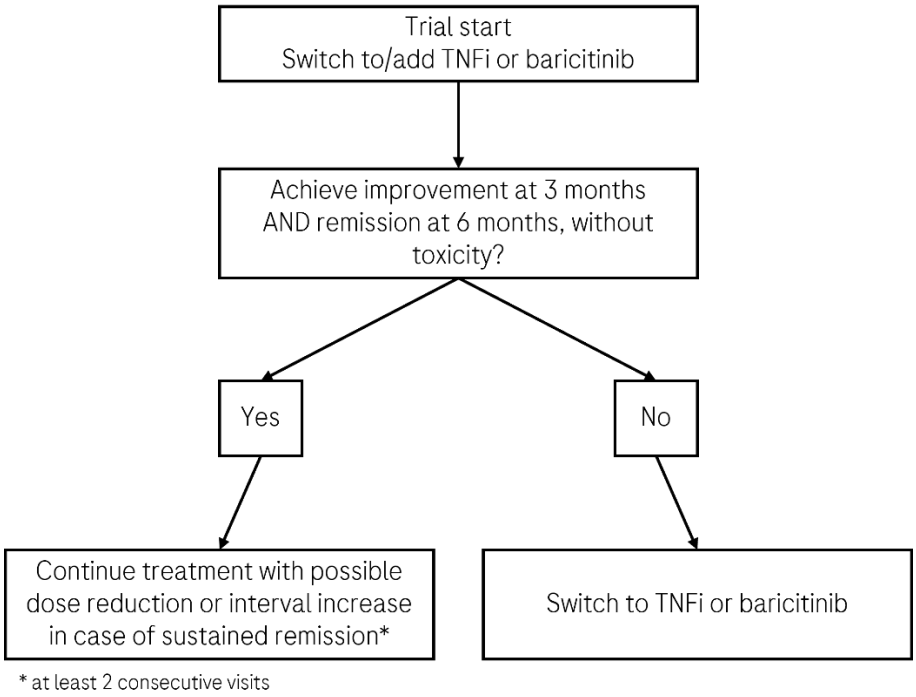
Procedures

Fifteen centres, of which fourteen were in the Netherlands and one in Belgium, who had all fully implemented a T2T strategy for RA before this study, were enrolled. Consecutive RA patients, not responding on or losing response on csDMARDs, were included after giving written informed consent. For all patients, a suggested treatment allocation to either to start with baricitinib or to start with a TNFi was provided by means of block-wise randomisation lists for each centre: final treatment decision was at the shared discretion of the attending physician and the patients, in compliance with the European Alliance of Associations for Rheumatology (EULAR) and national guidelines [180]. Both treatment strategies completely conformed to EULAR guidelines, the only difference between the two being the first treatment step after inadequate response to csDMARD is a choice that the 2019 EULAR guidelines did not yet distinguish between [188]. The 2022 EULAR guidelines update has included a caution around JAKi, prompting to consider them after risk assessment [191].

Patients were followed up over the course of 48 weeks with scheduled clinic visits at 0, 12, 24, 36, and 48 weeks and were encouraged to schedule visits if they experienced a disease flare or adverse events (AEs) in between scheduled visits. At each visit, disease activity guided therapeutic

adjustments were made as necessary in line with T2T principles, aiming to achieve clinical remission. Therapeutic adjustments included the option to taper or switch medication at the discretion of their attending physician. Patients were treated in accordance with the 2016 EULAR recommendations for RA treatment with synthetic and biological disease-modifying antirheumatic drugs including more recent updates thereof [180,188]. The protocol under investigation in PERFECTRA recommended the use of defined registered and reimbursed products.

Figure 1: Treatment Algorithm



TNFi, tumor necrosis factor inhibitor.

Figure 1 displays the study treatment algorithm. At baseline, patient characteristics including sex, age, and disease duration were collected. Clinical characteristics, including anti-cyclic citrullinated peptide (anti-CCP) antibody, rheumatoid factor status, comorbidities and medication use were also recorded at baseline and updated throughout the study period. During each visit, patients underwent full clinical assessments including laboratory testing of acute phase response (C-reactive protein (CRP)) and 28-joint count of tender and swollen joints (TJC and SJC). PROMs were completed online at 0, 4, 8, 12, 24, 36 and 48 weeks.

Measures

ACR50 (American College of Rheumatology 50) response criteria served as a primary endpoint, defined as a reduction of at least 50% in both the TJC and SJC and a reduction of at least 50% in three of the five following ACR core measures: physician global assessment of disease activity, patient global assessment of wellbeing, patient-reported pain, patient-reported disability, and CRP [192].

The physician global assessment of current disease activity was measured using a visual analogue scale (VAS) ranging from 0 (not active at all) to 100 (extremely active). Patient global assessments of pain in the past week (no pain at all – unbearable pain) , current wellbeing in the past week

(very well – very poor), and fatigue in the past week (not fatigued at all – extreme fatigue) were also assessed at every measurement point using 0-100 VASs. Disability was measured with the Rapid Health Assessment Questionnaire-II (RAPID3 HAQ) [193].

The composite 28-joint count Disease Activity Score with C-reactive protein (DAS28-CRP) was computed as a secondary endpoint [194]. The Clinical Disease Activity Index (CDAI) was also computed as some bDMARDs and tsDMARDs, including baricitinib, directly influence the C-reactive protein production. The CDAI shows disease activity results independent from the acute phase response [195].

Baseline and 48-week radiographs of hands and feet were scored according to the modified Sharp/van der Heijde method in random order by two trained readers independently [196]. 144 complete sets of radiographs were available, equally divided between both strategies. The readers were blinded to clinical information, chronological order, and strategies assigned. After confirming acceptable inter-reader reliability, the average score of the two readers was considered the Sharp/van der Heijde Score (SHS) and used for analysis.

Adverse Events (AE) and Serious Adverse Events (SAE), as defined by FDA, were obtained continuously during follow-up. Patients were asked to report any side effects experienced to their treating rheumatologist. Reporting and procedures were aligned with national guidelines [197].

Outcomes

The primary endpoint was non-inferiority, with subsequent superiority testing in case of non-inferiority with preservation of type 1 error rate [198], of the strategy of starting with baricitinib versus the comparator strategy to start with a TNFi, in terms of ACR50 response at 12 weeks. Secondary objectives included to compare the proportions of patients achieving DAS28-CRP < 2.6 at 12 weeks, changes in DAS28-CRP and CDAI scores and PROMs across the follow-up period, and radiological progression over 48 weeks. ACR response criteria and DAS28 both are composite outcome measures for RA. We chose to apply the ACR50 score as a realistic composite outcome measure since treatment decisions are usually driven by DAS28.

Drug survival

Switching from baricitinib to a TNFi or from TNFi to baricitinib was advised in case of no observed improvement or in case of intolerable side effects after 12 weeks or thereafter. Kaplan-Meier analysis with log-rank testing was performed to explore drug survival over 48 weeks in both strategies.

Sample size calculation

The required sample size for the primary non-inferiority analysis was based on an expected 35% of the patients in the baricitinib arm and 25% of patients in the TNFi arm obtaining an ACR50 response at 12 weeks. These estimates were obtained by adjusting the ACR50 response rates observed in the RA-BEAM trial [181] for differences in ACR50 response rates between clinical trial and clinical practice populations as described in previous studies [199,200]. The adjustment of ACR50 response rates was based on a random effects meta-analysis of the risk-difference of

obtaining ACR50 response in clinical practice versus clinical trial settings for biological medications in RA. To achieve 95% power for a non-inferiority test with a risk of type 1 error of 5%, 186 patients would need to be included to be sure that the lower limit of the 95% confidence interval for the difference in proportions of patients achieving ACR50 at 12 weeks would be above the prespecified non-inferiority limit of -12%. To account for dropout, we aimed to included 200 patients. Since PERFECTRA is a real world study, it was expected that the true difference between treatment arms might be lower than observed in RA-BEAM. To assess robustness of the sample size calculations against deviations from our initial expectations we tested a range of possible ACR50 response rates for the baricitinib arm ranging from 30-35% and the TNFi arm from 25-30%, with the difference between both arms ranging from 4-10%. The results showed that in all these scenarios power was ≥ 80% with the planned number of 200 included patients.

Populations and missing data

Primary endpoint analyses were performed for both the intention-to-treat (ITT) and per-protocol (PP) populations. The ITT population consisted of all subjects correctly included in the study, analysed based on assigned treatment. The PP population excluded all subjects that discontinued the study or had missing data for one or more assessments of the primary outcome at baseline or 12 weeks.

Assuming that any missing data occurred at random, missing values for the primary analyses were imputed for the ITT population using multiple imputation by chained equations (10 imputations with a maximum of 25 iterations) of the individual ACR50 components using predictive mean matching [201]. The imputation models were specified to include the individual component measures from which the ACR response criteria were calculated at baseline and 12 weeks as predictors along with baseline treatment group, sex, age, smoking status, disease duration, BMI, erosion, RF and anti-CCP positivity, and concomitant MTX use, based on previously established predictors of disease activity remission [202–204] All secondary analyses were performed using the available (non-imputed) data of all correctly, according to study protocol in- and exclusion criteria, included patients.

Primary effectiveness analyses

For the primary analysis, the proportion of patients achieving ACR50 at week 12 in baricitinib strategy arm were compared to the TNFi strategy arm, using the 95% Wilson score confidence interval for the difference in proportions using the Newcombe hybrid score [205,206].

Following previous studies, a fixed non-inferiority margin of 12% was adopted for this study [181,207,208]. If the 95% confidence interval for the difference in proportions of patients achieving ACR50 at 12 weeks (TNFi – Baricitinib) lies entirely to the right of -12%, baricitinib will be declared non-inferior. The prespecified NI margin of 12% was based on previous head-to-head trials in RA, including the RA-BEAM baricitinib study [181].

Secondary analyses

The difference in the proportion of patients achieving DAS28-CRP <2.6 was also compared using the 95% Wilson score confidence interval method for the difference in proportions in those

patients with available data at both baseline and 12 weeks. Continuous secondary endpoints were analysed using linear mixed-effects (LME) models with the endpoint as the dependent variable and time, treatment group and their interaction as fixed effects and random effects for patient intercepts and slopes over time. All LME models used restricted maximum likelihood estimation, and the covariance structure was set to compound symmetry as unstructured or autoregressive structures did not provide significantly better fit according to likelihood ratio tests for disease activity and PROMs.

Radiological joint damage of hands and feet scores (SHS) were analysed by performing a Mann-Whitney U test for difference in progression scores at 48 weeks due to their non-normal distribution (zero-inflated or positively skewed) and were visualised in cumulative probability plots [209,210].

Safety analyses

Safety was evaluated by tabulations of AE/SAE and presented with descriptive statistics for each treatment group.

Patient involvement

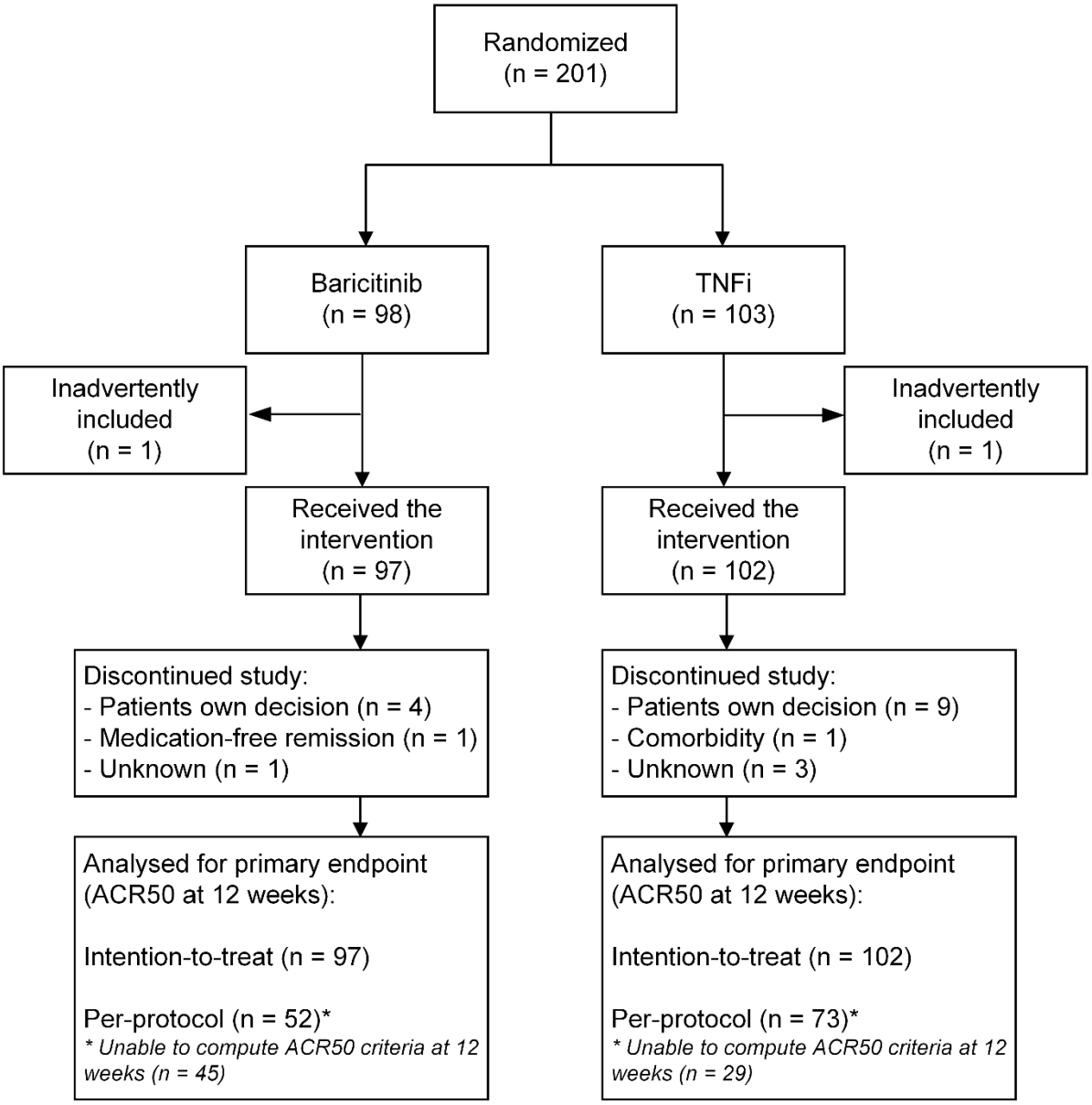
Patients were involved in the design of PERFECTRA. The Dutch patient association 'Nationale Vereniging ReumaZorg Nederland (RZN)' was involved with design of the study protocol and were consulted on design and feasibility.

Results

Patients

Inclusion started on September 25th, 2019, and the last patient was included on February 2nd, 2022. Last patient out took place one on April 4th, 2023. In total, 201 patients were included in the study of which 199 patients received a first dose. Figure 2 displays the CONSORT flow chart [211,212]. After randomization, 97 patients were assigned and started baricitinib strategy where 102 patients were assigned to and started TNFi strategy. Within the TNFi strategy, 64% of patients started on adalimumab, 33% used etanercept, the rest of the group used golimumab or infliximab.

Figure 2: Study Flow Diagram



TNFi, tumor necrosis factor inhibitor

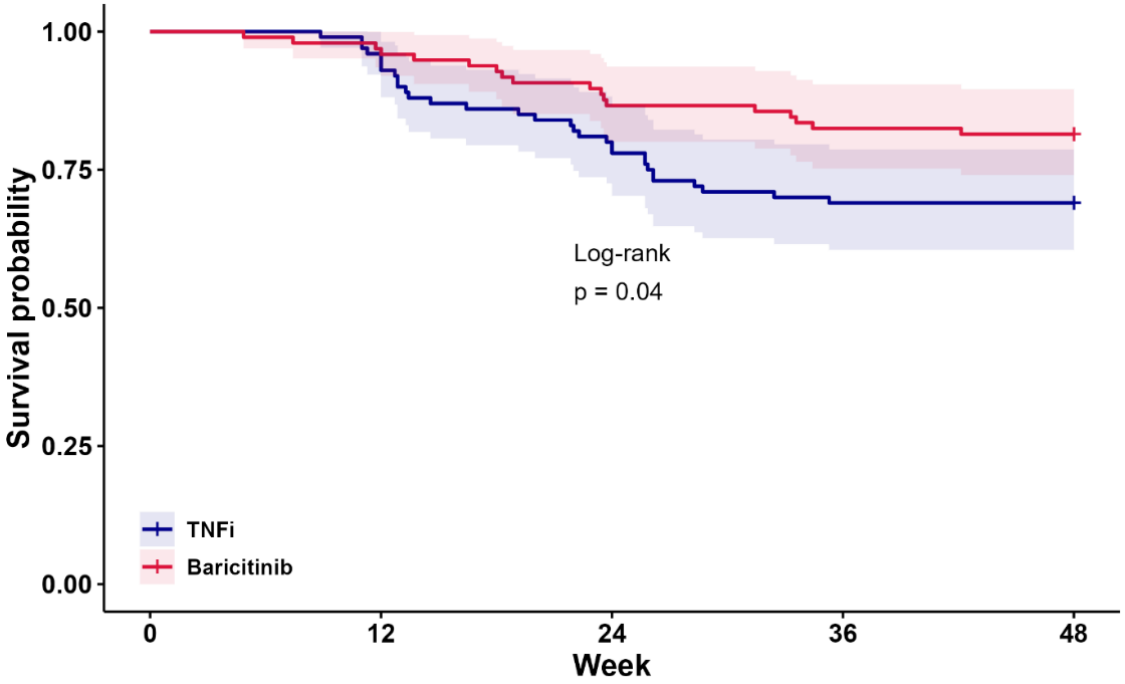
Table 1 shows baseline characteristics for the study sample. Baseline characteristics were similar in both treatment groups. About two thirds of the patients were female with an average disease duration of 2 years since diagnosis with RA. Disease activity as measured by DAS28-CRP scores, on average, match with moderate disease activity [181,186]. Concomitant MTX use and changes in dose of MTX are displayed in more detail in the supplemental material.

Table 1 Baseline Characteristics

	TNFi strategy	Baricitinib strategy
	N=102	N=97
Age (years) , mean(SD)	55.2 (13.4)	54.8 (12.0)
Female, n (%)	68 (66.7%)	62 (63.9%)
Smoking, n (%)		
Never	38 (37.3%)	37 (38.9%)
Stopped	39 (38.2%)	36 (37.9%)
Yes	25 (24.5%)	22 (23.2%)
BMI, mean (SD) kg/m^2	27.4 (4.93)	26.5 (5.03)
Disease duration (years), median (IQR)	2.00 [1.00;3.00]	2.00 [1.00;3.00]
Erosions state, n (%)		
No	66 (64.7%)	69 (71.1%)
Unknown	19 (18.6%)	16 (16.3%)
Yes	17 (16.7%)	12 (12.4
CV		
No	74 (72.5%)	76 (78.4%)
Yes	28 (27.5%)	21 (21.6%)
Rheumatoid factor positive, n (%)	69 (67.6%)	70 (72.2%)
ACCP+, n(%)	65 (63.7%)	70 (72.2%)
Concomitant MTX, n(%)	69 (67.6%)	62 (63.9)
MTX dose, mg per week, mean (SD)	19.9 (5.06)	20.48 (5.42)
Glucocorticosteroid, n (average dose (mg)/day)	16 (6.3)	28 (7.7)
DAS28-ESR, mean (SD)	4.43 (1.06)	4.41 (1.14)
DAS28-CRP, mean (SD)	4.17 (1.03)	4.08 (1.05)
TJC, median (IQR)	4.00 [2.00;7.00]	4.00 [2.00;7.00]
SJC, median (IQR)	3.00 [1.00;5.00]	3.00 [2.00;4.00]
ESR, mm/h, mean (SD)	24.0 (19.5)	25.1 (22.1)
CRP, mg/L mean (SD)	13.7 (19.1)	12.3 (17.5)
physician global, mean (SD)	50.0 (21.3)	51.9 (16.9)
VAS Wellbeing, mean (SD)	61.1 (21.5)	54.1 (22.5)
VAS pain, mean (SD)	61.5 (24.0)	55.6 (25.0)
RAPID 3 HAQ score, mean (SD)	12.2 (5.66)	10.9 (6.38)
CDAI, mean (SD)	19.9 (8.68)	19.6 (8.63)

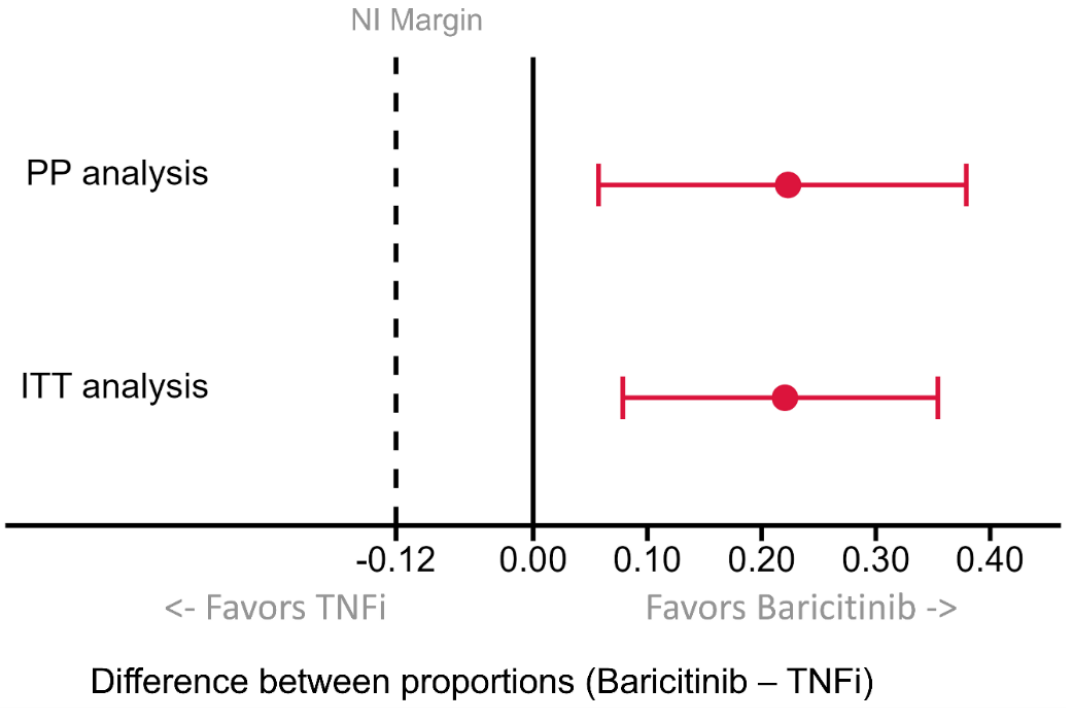
TNFi: Tumor necrosis factor inhibitors, BMI: Body Mass Index, CV: increased CardioVascular risk as reported by attending physician, ACCP: anti-cyclic citrullinated peptide antibody, MTX: methotrexate, TJC: tender joint count, SJC: swollen joint count, ESR: erythrocyte sedimentation rate, CRP: C-reactive protein, PG: physician global, VAS: visual analogue scale, Rapid 3 HAQ: Health Assessment Questionnaire, CDAI: Clinical Disease Activity Index

Figure 3: Drug Survival in TNFi Strategy and Baricitinib Strategy



TNFi, tumor necrosis factor inhibitors

Figure 4: Difference between proportions in achieving American College of Rheumatology 50 at 12 weeks (baricitinib – TNFi).



PP, per protocol; ITT, intention-to-treat; NI margin, non-inferiority margin; TNFi, tumor necrosis factor inhibitors.

Figure 3 displays the survival probability of baricitinib and TNFi in both strategies by Kaplan Meier plots. Throughout the entire study period, approximately 70% of patients in TNFi strategy remained on their first treatment versus around 80% for baricitinib (p=0.04). During the study twenty-seven patients in the TNFi first strategy switched to baricitinib, while 4 switched to an interleukin 6 inhibitor (IL-6i). In the baricitinib strategy, 15 patients switched to a TNFi, while 2 switched to an IL6i, and one patient stopped baricitinib because of a stable clinical remission. At 12 weeks, only 7 patients in TNFi strategy and 4 patients in baricitinib strategy switched at that point.

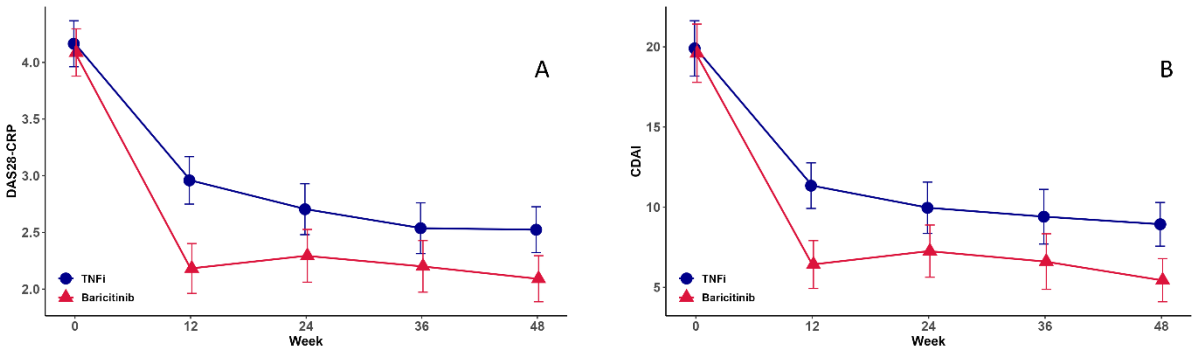
Primary endpoint

Figure 4 shows the difference in proportions of ACR50 achievement at 12 weeks between the baricitinib strategy and TNFi first strategy (PP: Δ22%, 95% Wilson CI: 5.7% - 38%, ITT: Δ22%, 95% Wilson CI: 7.8% - 35%). In the PP population, 23/52 (44%) of baricitinib strategy patients reached ACR50 response at 12 weeks (ITT population: 42%), compared to 16/73 (22%) in the TNFi strategy group (ITT population: 20%). At 12 weeks, the lower bound of the 95% Wilson score interval for the difference in proportions of patients meeting the ACR50 response was to above the -12% non-inferiority margin and the right of zero in both the PP and ITT analysis. Hence, baricitinib was found to be not only non-inferior but also statistically superior to TNFi in the analysis of the primary endpoint in both the per-protocol and ITT analyses. Mean scores in individual ACR50 components can be found in supplemental material.

Secondary endpoints

Of the patients in the baricitinib strategy, 65/87 (75%) (ITT population: 75%) reached DAS28-CRP <2.6 at 12 weeks. This was significantly more than the 45/97 (46%) (ITT population: 46%) in the TNFi strategy group (PP:Δ28%, 95% Wilson CI: 14% - 41%, ITT: Δ28%, 95% Wilson CI: 13% - 41%). Mean DAS28-CRP scores in both study groups showed a strong decline from baseline to 12 weeks and a gradual further decrease up to week 48, as shown in Figure 5, panel A. There was a more rapid decline in DAS28-CRP scores in the baricitinib strategy, compared to the TNFi strategy. Throughout the study period, the estimated marginal means remained lower in the baricitinib strategy. CDAI scores, that do not include an acute phase reactant, showed a comparable pattern (Figure 5, panel B). The individual DAS28 components all showed a comparable pattern in favour of baricitinib strategy (see supplemental material).

Figure 5: Estimated marginal means in 28-joint count Disease Activity Score with C-reactive protein (DAS28-CRP) and Clinical Disease Activity Index (CDAI) at weeks 0, 12, 24, 36, and 48.



Error bars represent 95% WALD Confidence interval. TNFI, tumor necrosis factor inhibitors.

All PROMs showed a clear improvement for both strategies, generally in favour of the baricitinib strategy. Differences in PROMs scores persisted over the full length of the study period. Plots for wellbeing, disease activity, pain, disability, and fatigue can be found in the supplemental material along with a table containing fixed effects estimates from the LME analyses.

Radiological damage was limited in both treatment strategies with the majority of patients showing no progression over 48 weeks. No significant difference was found in progression scores between the groups (p=0.246), although a small numerical benefit was noticeable for baricitinib. The SHS cumulative probability plot and median radiographic progression scores can be found in the supplemental material.

Safety

The majority of SAEs were, according to the prescribing physician, not related to baricitinib/TNFi with the exception of three reported cases of infection. According to the prescriber, all three infection cases were possibly attributed to baricitinib. One reported case of gastrointestinal (GI) complication, ic “diaphragmatic hernia after gastric bypass surgery”, could be attributed to TNFi (etanercept) according to the physician. All other SAE were, according to the physician not attributable to used DMARDs. The overall incidence of AEs and their nature was comparable across the two strategies (see Table 2; Full AE table can be found in the supplemental material).

Table 2 Safety

	Group analyses					Present treatment analyses		
	Baricitinib strategy	% of total	TNFi strategy	% of total	Total	Patients on Baricitinib	Patients on TNFi	Total
SAE:	6	6.2%	5	4.9%	11	6	4	10
"cancer"	0	0.0%	2	2.0%	2	0	2	2
"GI complication"	0	0.0%	2	2.0%	2	0	1	1
"infections"	3	3.1%	0	0.0%	3	3	0	3
"total knee arthroplasty"	1	1.0%	0	0.0%	1	1	0	1
"MI"	1	1.0%	0	0.0%	1	1	0	1
"fracture"	0	0.0%	1	1.0%	1	0	1	1
"cerebral concussion"	1	1.0%	0	0.0%	1	1	0	1
AE	111		144		255	115	129	244

Present treatment analyses: (s)AE event developed while on either baricitinib of TNFi. N Baricitinib strategy group: 97, N TNFi strategy group: 102, number of patients receiving baricitinib at any point during study: 124, number of patients receiving TNFi at any point during study: 117, SAE: Serious Adverse Events, GI: Gastro Intestinal, MI: myocardial infraction, AE: adverse event, TNFi: Tumor necrosis factor inhibitors

Discussion

The PERFECTRA study addressed the question of where we stand with a tsDMARD first versus a TNFi first treatment strategy for RA patients failing to achieve DAS28-CRP <2.6 on csDMARDs [213]. This real-world study in a T2T setting applying EULAR guidelines [191] showed that after failing csDMARDs baricitinib is non-inferior and superior as compared to TNFi strategy with respect to clinical effectiveness, PROMs, and drug survival. This study was not powered to address neither comparative safety nor radiological damage between baricitinib and TNFi. We did not observe adverse events that have not previously reported with either therapy. When looking at the DAS28-CRP and CDAI graphs in figure 5 (and the individual component score in the supplemental material) one should be aware that the real distinguishing moment is at 12 weeks. In line with the T2T principles, non-responders were encouraged to switch after 12 weeks, which could cause convergence between strategies. . Progression of radiological damage was comparable for both treatment strategies, with only very few patients showing any progression over 48 weeks of treatment. Such zero-inflated data presents extensive modelling challenges. It also should be noted that 144 complete radiographs sets were found indicating some patients had missing sets.

A preceding pivotal RCT compared baricitinib with adalimumab, RA BEAM [181]. PERFECTRA suggests that in a T2 setting, there is more likelihood of a response to baricitinib than adalimumab and possibly other TNFi, which is in line with findings from RA BEAM. In PERFECTRA, the TNFi strategy led to 46% of patients being in remission after 12 weeks, this number adds up to 75% for the baricitinib strategy. These results seem high as compared to results RA BEAM, however, background of patients was vastly different between PERFECTRA and RA BEAM's patient populations. RA BEAM has an average prior disease duration of 10 years, whereas PERFECTRA average duration was 2 years. Pre-study treatment also differed significantly. PERFECTRA study patients where previously already treated with T2T principles, this was not warranted in RA BEAM. Another major difference is that in RA BEAM all included patients had to have erosions. Composite disease activity scores are not only driven by inflammation alone but also by damage, this is illustrated by the average DAS score of >6 in RA BEAM at baseline. For all these reasons, achieving DAS28-CRP <2.6 in long-standing erosive RA, like in the RA-BEAM study population, is more difficult than in the early RA population, like in the PERFECTRA study. Remission rates in the PERFECTRA study are in line with results from other T2T studies [214–218].

Inhibitors of Janus Kinase other than baricitinib are also approved for the treatment of RA. For tofacitinib [182,219], filgotinib [183,220],and upadacitinib [184,221,222] formal RCTs were performed comparable to the RA-BEAM study in non-T2T setting, showing comparable or even more favourable results for JAKi vs. TNFi. Whether the additional evidence by the present study for baricitinib in the T2T setting can be extended to the other JAKi has yet to be determined. Not only are multiple JAKi available, but also many TNFi are approved for the treatment of RA. Most formal RCTs, as cited earlier, used adalimumab as their comparator. PERFECTRA allowed all approved TNFi. The distribution of prescribed TNFi in our study reflects the market in the Netherlands meaning etanercept and adalimumab were predominantly used, combining to approximately 95%. We did not find any indication for different responses between TNFi used in

the study, indicating that the conclusion of superior effectiveness of baricitinib strategy holds true for TNFi used.

The PERFECTRA study was designed for the currently relevant setting, where T2T is fully implemented. The rheumatological care preceding the inclusion of patients was according to this recommendation. The study was adequately powered for the primary outcome (ACR response at 12 weeks). For other clinical outcomes and PROMs over the study period of 48 weeks consistent relevant difference between both strategies were found in favour of Baricitinib. For a full picture on safety and joint damage a larger population and a longer study period would be necessary.

PERFECTRA was performed during the COVID19 pandemic. To the best of our knowledge, the consecutive inclusion and follow up of patients was, although challenged, not jeopardized. The only exception was capturing ESR, since due to COVID19 measuring ESR was because of an attempt to reduce the contamination risk, skipped in many centres.

One might question if the direct effect of baricitinib on CRP is responsible for the observed superior efficacy. Therefore, we also reported the CDAI scores, which showed similar results with respect to disease activity, independent from the acute phase reaction. The individual disease activity components also showed highly comparable results.

Oral glucocorticoid (<10 mg/day) and NSAID use were limited. Switching, and stopping throughout the study were not advised but always at the discretion of the attending physician. Consequently, this might not always be collected in the e-CRF resulting in limiting analysing ability of these factors.

Not all TNFi strategy patients used MTX at start and throughout the study. In the real world, MTX is increasingly perceived as poorly tolerated. Even though TNFi should preferably be used in combination with MTX, this is not always the case in practice. This study shows the comparative effectiveness of TNFi vs. JAKi in daily clinical practice, including all nuances, contingencies, and issues. Real World adherence to MTX due to intolerance or other factors should be seriously considered when comparing treatment (sequences). MTX use add baseline and during the study was comparable in both groups

DAS28-CRP, as used in PERFECTRA can pose some challenges, it should be noted that DAS28-CRP scores tend to be slightly lower than DAS28-ESR scores [223,224]. Having said this, DAS28-CRP is frequently chosen as the proxy for measuring disease activity in daily clinical practice.

We designed PERFECTRA as a randomized, open label controlled real-world strategy study. Although we acknowledge that in theory the compromises that have to be made in pragmatic studies may result in conclusions that are methodologically less rigorous than those resulting from formal RCTs. Randomization by physician election could raise the question of physician bias. In order to mitigate, baseline characteristics were very carefully analysed and deemed appropriate, PERFECTRA showed nicely comparable groups at baseline. In terms of robustness and generalizability, pragmatic studies can add knowledge to pivotal trials. Generalizability of study

results should always be carefully done. We think that the PERFECTRA study provides relevant information especially for the T2T setting of early-RA patients after failure of cs-DMARDs.

To this date, for good reasons, medical science and regular authorities heavily lean on formal RCTs. However, If conclusions of these pivotal trials can be successfully implemented in Real World, where no preselection due to disease characteristics, age, comedication comorbidities and the challenges by the healthcare system are relevant challenges. PERFECTRA, addresses in a real world T2T setting the clinical questions whether to start a JAKi (baricitinib) or any TNFi after csDMARDs failure. On the one hand, a real-world open design may theoretically be methodologically less robust than a formal randomized and double-blinded RCT, the generalizability and applicability of this solid pragmatic real-life study complement the formal RCTs by demonstrating the effectiveness in a real-world setting in daily clinical practice, something which formal RCTs can lack. PERFECTRA not only confirms previous suggestions by preceding RCTs and registry data that baricitinib, and probably JAKi in general, have increased efficacy as compared with TNFi (adalimumab), but illustrates that, in combination with a fully implemented T2T approach to the target of DAS28-CRP <2.6, the decision to start with baricitinib after failure of csDMARDS is a valid option for patients. Results from a real world setting not only complement formal RCTs but can also assist in identifying proper application in specific situations for physicians.

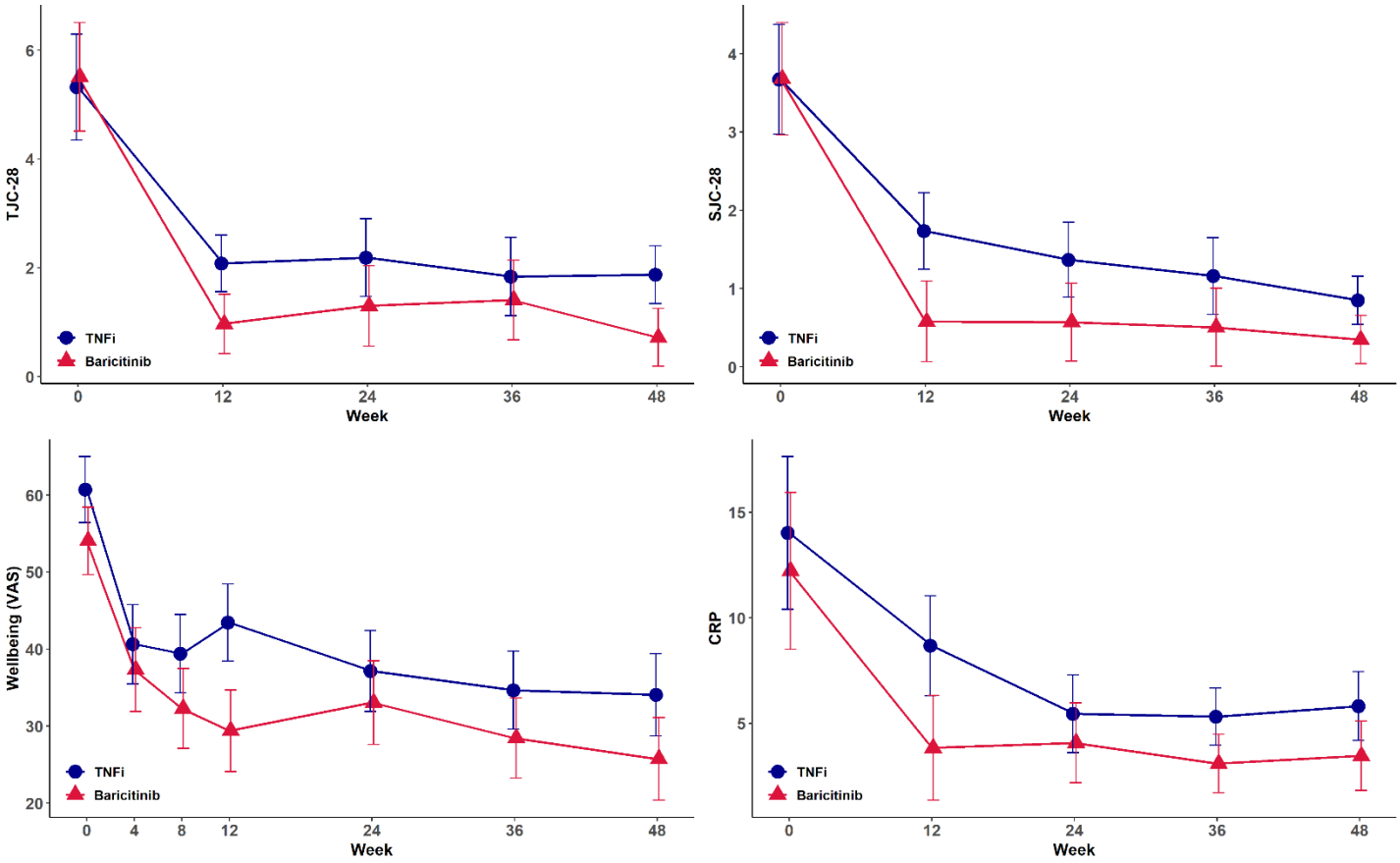
Although not powered for safety, no unexpected safety signals were seen in this relatively small group of patients. Prescribers have to be aware that cardiovascular and malignant serious adverse events are more frequently reported in JAKis than TNFis. Obviously, this has to be considered carefully in risk-benefit discussions with any individual patient. Nonetheless, there are well known safety warnings for both JAKi [225,226] and TNFi [227]. For treatment with JAKi, full history and physical examination and testing, among other things as recommended, and outlined fully in Nash et al. [189]. Rheumatologists should be aware and are educated to take into account the individual patient's safety profile as well as the patient's preferences in order to minimalize the risk of complications and to balance these risks with the expected efficacy. Rheumatologists learned to effectively manage the complication risks of TNFi (among others, cardiovascular events, congestive heart failure and infections, especially tuberculosis) [228,229]. As of more recently, they manage the warnings for JAKi (among others, herpes zoster, major arterial cardiovascular events (MACE) and thromboembolism).

Conclusions

In the setting of real world T2T treatment for RA, as advised by professional societies like EULAR, after failure on csDMARDs, PERFECTRA suggests that the strategy to start baricitinib is a feasible alternative to starting with TNFi with respect to disease activity and PROMs. Baricitinib also showed beneficial drug survival compared to that of TNFi. The limited number of patients and study duration do not allow for conclusions on differences in safety and radiological damage which need to be established in larger and longer studies.

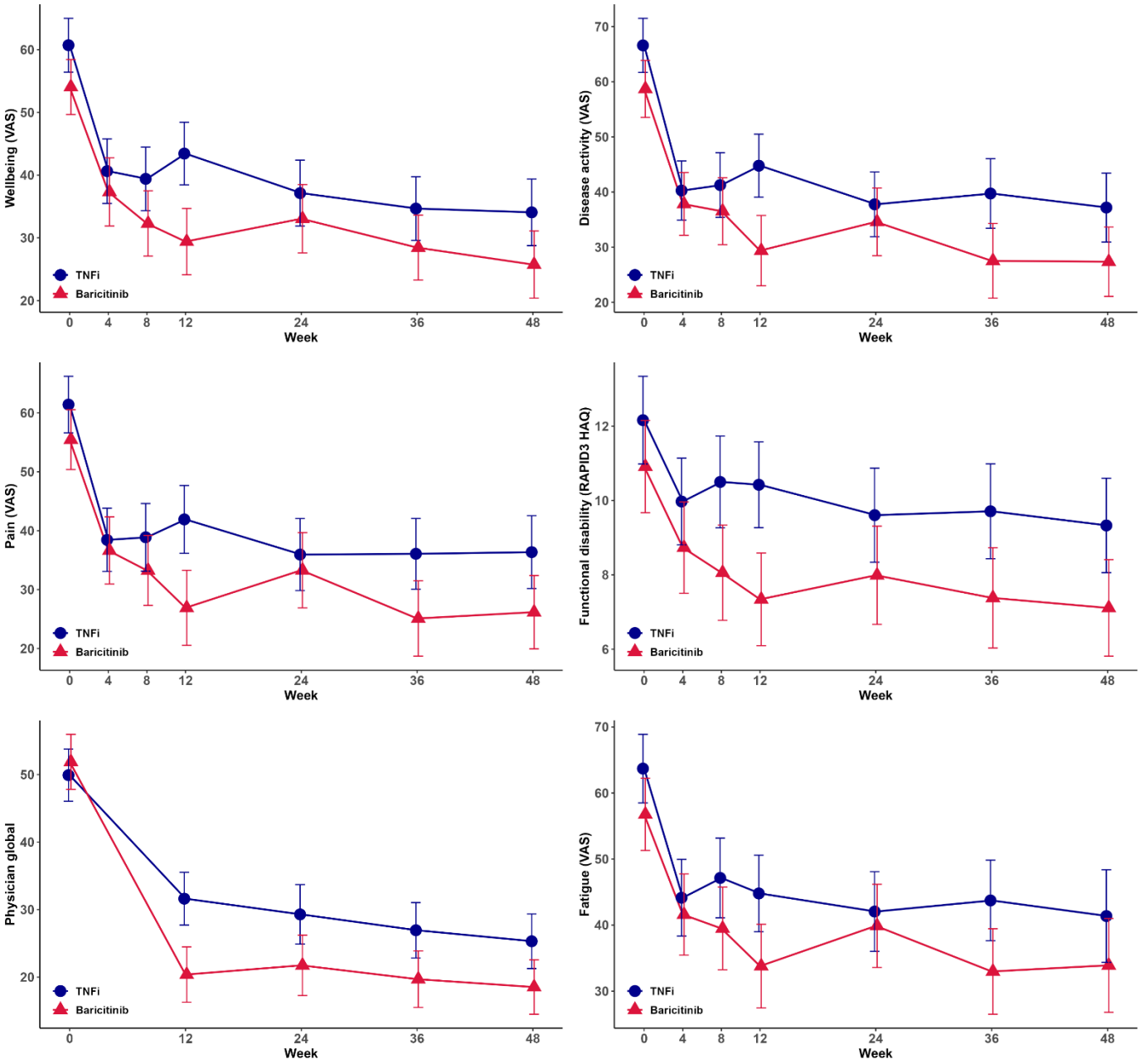
Supplemental Material

Supplement Figure 1: Estimated Marginal Means in individual DAS28-CRP components at weeks 0, 12, 24, 36, and 48



Error bars represent 95% Wald CI. DAS28-CRP: Disease Activity Score with 28-joint count C-reactive protein, TNFi: Tumor necrosis factor inhibitors, VAS: visual analog scale, TJC: tender joint count, SJC: swollen joint count

Supplement Figure 2 Estimated Marginal Means in VAS Wellbeing, VAS Disease Activity, VAS Pain, functional disability with HAQ, physician global and VAS Fatigue at weeks 0, 12, 24, 36, and 48



Error bars represent 95% Wald CI. DAS28-CRP: Disease Activity Score with 28-joint count C-reactive protein, TNFi: Tumor necrosis factor inhibitors, VAS: visual analog scale, TJC: tender joint count, SJC: swollen joint count, RAPID3 HAQ: Rapid Health Assessment Questionnaire-II

Supplement Table 1 Fixed effects estimates of mixed model analysis for secondary endpoints outcome measures

	Intercept	Group		Time		Group x Time	
Measure	value (std. error)	value (std. error)	p-value	value (std. error)	p-value	value (std. error)	p-value
DAS28-CRP	3.7	-0.3 (0.14)	0.00	0.0 (0.00)	0.000	0.0 (0.00)	0.47
CDAI	16.9	-2.0 (1.05)	0.05	-0.2 (0.03)	0.318	0.0 (0.00)	0.32
CRP	11.8	-2.7 (2.02)	0.00	-0.2 (0.040)	0.00	0.0 (0.05)	0.92
ESR	22.6	-2.6 (2.78)	0.35	-0.2 (0.04)	0.00	0.0 (0.06)	0.93
SJC	3.05	-0.47 (0.32)	0.13	-0.05 (0.00)	0.00	-0.01 (0.01)	0.57
TJC	4.2	-0.2 (0.53)	0.65	-0.1 (0.01)	0.00	0.0 (0.01)	0.24
Physician Global	44.0	-4 (2.55)	0.15	0.0 (0.05)	0.00	0.0 (0.08)	0.20
Wellbeing	49.0	-7.0 (1.93)	0.00	0.0 (0.06)	0.00	0.0 (0.08)	0.89
RAPID3 HAQ	11.1	-1.7 (0.78)	0.03	0.0 (0.01)	0.0001	0.0 (0.02)	0.42
Pain	48.0	-5.0 (3.07)	0.10	0.0 (0.06)	0.00	0.0 (0.09)	0.32
Fatigue	53.0	-6.0 (3.50)	0.10	0.0 (0.10)	0.000	0.0 (0.10)	0.60

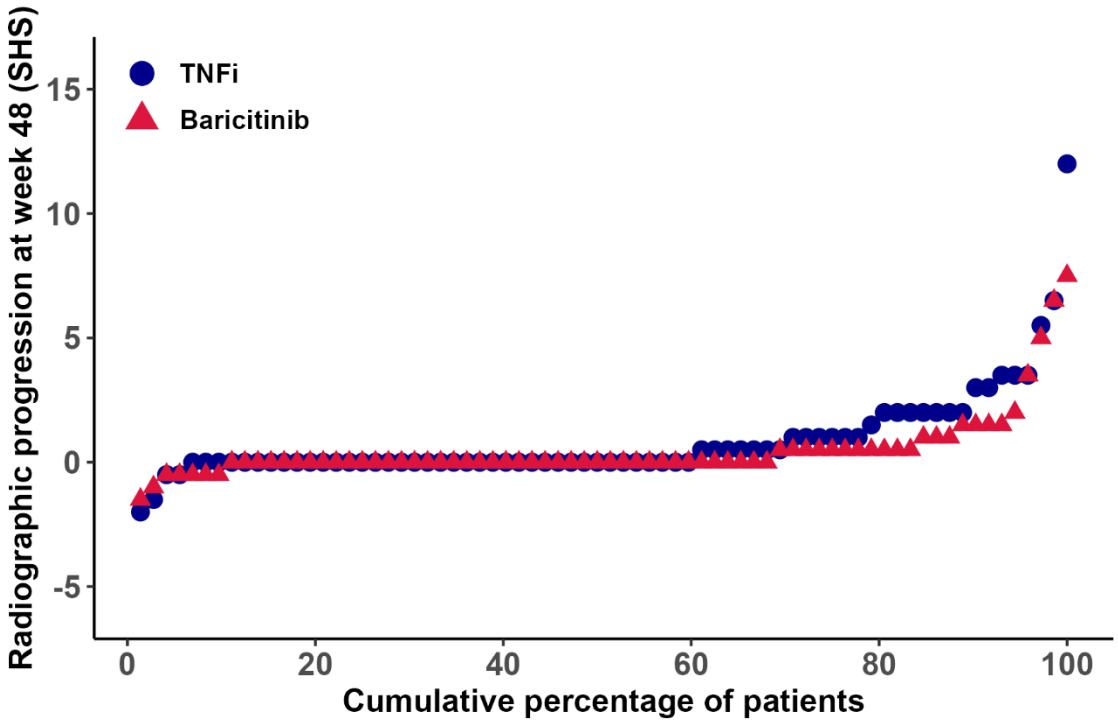
DAS28-CRP: Disease Activity Score with 28-joint count C-reactive protein, CDAI: clinical disease activity index, CRP: C-reactive protein, ESR: erythrocyte sedimentation rate, SJC: swollen joint count, TJC: tender joint count, RAPID3 HAQ: Rapid Health Assessment Questionnaire-II

Supplement Table 2: Median SHS scores at baseline and week 48 with 95% Wald CI

	TNFi strategy	Baricitinib strategy	P-value overall
Week 0	0.00 (0.00; 2.00)	0.00 (0.00; 1.12)	0.407
Week 48	1.00 (0.00; 3.12)	0.00 (0.00; 1.62)	0.099
Progression	0.00 (0.00; 1.00)	0.00 (0.00; 0.50)	0.121

SHS: Sharp/van der Heijde Scores, TNFi: Tumor necrosis factor inhibitors

Supplement Figure 3: SHS cumulative probability plot



SHS: Sharp/van der Heijde Scores, TNFi: Tumor necrosis factor inhibitors

Supplement Table 3: SAE and AE full overview

		Baricitinib first	TNFi first	While on baricitinib	While on TNFi
		N			
SAE	tibiaplateau fracture left after trauma	0	1	0	1
	myocard infacrt	1	0	1	0
	hernia diafragmatica after gastric bypass$^{\beta}$	0	1	0	0
	beklemde navelbreuk	0	1	0	1
	bilateral pneumonia$^{\theta}$	1	0	1	0
	infection of urinary tract$^{\theta}$	1	0	1	0
	conjunctivitis$^{\theta}$	1	0	1	0
	total knee arthroplasty	1	0	1	0
	gemetastaseerd cholangiocarcinoom	0	1	0	1
	esophageal carcinoma	0	1	0	1
	cerebral concussion	1	0	1	0
airways	dyspnoea	0	1	0	1
blood	glucose$^{\delta\ (2)\ \theta\ \alpha}$	3	1	1	3
	increase of cholesterol$^{\theta\ (2)\ \phi}$	3	0	3	0
	increase of lipids$^{\theta}$	1	0	1	0
	leucopenia$^{\alpha}$	0	1	0	1
carcinoma	carcinoma	0	1	0	1
cold	cold$^{\delta\ \lambda\ \alpha}$	2	1	0	3
cough	cough$^{\theta\ \alpha}$	1	1	1	1
cramp	cramp	0	1	0	1
disordered appetite	disordered appetite$^{\alpha\ (2)\ \varsigma}$	4	0	4	0
dizziness	dizziness$^{\delta\ \mu}$	1	3	0	4
edema	edema$^{\xi}$	1	0	1	0
fatigue	fatigue$^{\delta\ \theta\ (3)\ \alpha\ (3)}$	6	2	6	2
fever	fever$^{\mu}$	2	0	1	1
fractures	fractures	2	1	1	2
gaining	gaining weight$^{\delta\ \theta\ (4)}$	4	1	4	1
GI tract	diarrhea$^{\theta\ \alpha}$	3	0	3	0
	GI complaints$^{\theta\ (2)\ \varepsilon\ (2)\ \nu\ \alpha\ (2)}$	4	4	6	1
	naussea$^{\delta\ (10)\ \theta\ (4)\ \rho\ \alpha\ (11)}$	5	24	10	19
head	dry eyes$^{\delta}$	0	1	0	1

	dry nose[δ]	0	1	0	1
	runny nose[δ]	0	1	0	1
	tinnitus[δ]	1	2	1	2
hypertension	hypertension [δ (2) θ (2) χ α]	3	3	3	2
Infection	airways[π (2) θ γ (3) α]	1	7	1	7
	bursitis[γ α]	0	2	0	0
	cold sore[θ γ σ α]	1	3	1	2
	COVID-19[γ]	4	2	2	3
	eye[θ γ α]	1	2	2	1
	flu[μ]	1	2	0	3
	fungal[δ]	1	1	1	1
	prostate	1	0	1	0
	shingles[θ]	1	0	1	0
	sinusitis[θ]	0	1	1	0
	skin[θ β α]	0	3	2	1
	throat[δ]	0	1	0	1
	tooth	0	1	0	1
	urinary tract[δ (2) θ (6) γ (2)]	5	7	6	6
	viral[θ]	1	0	1	0
inflammation	tendon	0	2	0	2
	throat[β (2)]	0	2	0	2
insomnia	insomnia[δ]	1	1	1	1
kidney	renal dysfunction[ζ θ]	1	1	2	0
less blood loss during period	less blood loss during period	1	0	1	0
liver	increased liver values[θ ε τ α (8)]	2	10	4	6
Malaise	malaise[δ]	0	1	0	1
Mouth	burning sensation[θ]	1	0	1	0
	change of taste[δ θ (2)]	2	1	2	1
	mouth sore[θ]	1	0	1	0
	mucus[α]	0	1	0	1
	sensitive skin[θ]	1	0	1	0
Other	Body feels warm after Enbrel injection[γ]	0	1	0	1
	dental problems[α]	0	1	0	1
	Dupuytren	1	0	0	1
	fall	1	0	1	0
	feeling of pressure on the chest[α]	0	1	1	0

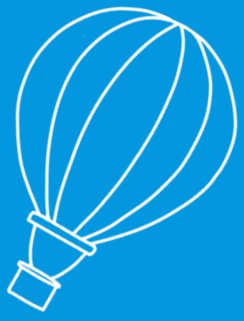
	Hernia cicatricalis	1	0	1	0
	hernia with ulcus	1	0	0	1
	HNP	1	0	1	0
	hot flashes$^{\theta\ (2)\ \gamma}$	2	1	2	1
	increase activity glandula parotis right	0	1	0	1
	increase PSA value	0	1	0	1
	Macrocytaire anemia$^{\theta}$	1	0	1	0
	mortonse neurose right foot	0	1	0	1
	Not feeling well for a few hours$^{\delta}$	0	1	0	1
	numbness and tingling sensation of hands and feet$^{\eta}$	0	1	0	1
	shaky$^{\alpha\ \delta}$	2	0	0	2
	swollen hands and feet during hot summer	1	0	1	0
	trombosflebitis	0	1	0	1
Pain	chest$^{\lambda}$	0	2	0	2
	head$^{\delta\ \theta\ (5)\ \gamma\ \alpha}$	9	6	11	4
	leg$^{\theta\ \gamma}$	1	1	1	1
	stomach$^{\alpha}$	0	1	0	1
	throat$^{\theta}$	0	2	1	1
perspiration	perspiration$^{\mu}$	1	0	0	1
restless	restless$^{\iota}$	1	0	1	0
skin	bruise$^{\delta\ \theta\ \gamma\ \alpha\ \iota}$	3	3	4	2
	cellulitis$^{\theta}$	1	0	1	0
	rash/itch/eczema$^{\delta\ (5)\ \theta\ (7)\ \gamma\ (4)\ \varepsilon\ \alpha\ \tau\ (2)\ \kappa}$	8	17	7	15
	wound	0	1	0	1
aggravation of complaints	omarthrose right$^{\iota}$	1	0	1	0
	rheumatoide artritis$^{\theta}$	2	0	2	0

Medication use related to the (S)AE's according to physicians is indicated by footnotes: α Methotrexaat, β Benepali, γ Etanercept, δ Adalimumab, ε Hydroxychloroquine, ζ Arthrotec, η Pravastatine, ϑ baricitinib, ι Prednisolon, κ Tocilizumab, λ Leflunomide, μ Infliximab, ν Meloxicam, ξ Amlodipine, π amgevita, ρ Celecoxib, ς Sulfasalazine, σ Filgotinib, τ Sarilumab, φ rosuvastatine, χ hydrochloorthiazide. The number in brackets following the footnote indicates the amount of AE (>1) that were thought to be related to the medication used according to physicians, TNFi: tumor necrosis factor inhibitors


Chapter 7







Chapter 7 – General Discussion

In all healthcare systems, decisions must be made continuously. First, the financial and organizational structure determine the effectiveness and efficiency of the designed system. The design of a healthcare system is the prerogative of the government. In the Western world, we hope that the structure and financial basis of a healthcare system is based on knowledge and data in order to ultimately be as optimal as possible in the interest of the population and its health status. Within an established healthcare system, health policy makers are on a continuous basis in need of data. First and foremost, management and financial information, but information on proposed innovations and adaptations as well. The concept to increase effectiveness of healthcare at affordable costs is widely accepted.

It needs not to be discussed that healthcare suffers from complexity. From care to cure, for the broad spectrum of acute and chronic health problems, including all diseases and conditions, has to be positioned in any healthcare system. We must be aware that fundamental organisational choices between a centrally organised and financed health care system and a decentral organised system have consequences for health policy makers and their work.

Management information is generally good and sometimes abundantly available, but usable outcome information remains a major challenge. The field of health economics has the tools to perform retrospective or prospective analyses and to provide information for policy makers. It is clear that a decentralised healthcare system with many legal entities as executors will have more difficulty in obtaining reliable data on outcomes and costs, because each legal entity can be unique in these respects. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) has provided a comprehensive approach to Health Technology Assessment (HTA) evaluating multiple dimensions to assess added value of interventions.

The ISPOR HTA Dimensions framework concerns clinical effectiveness, safety, economic implications, ethical, social, cultural and legal issues, organisational and environmental aspects, and lastly, patient and population impact [39,40]. In this dissertation I have conducted a number of studies that are in line with ISPOR Core HTA model, more specifically relating to current use, clinical effectiveness and cost & economic considerations. The various studies are examples on the one hand of the healthcare domains (rheumatology) and systems (predominantly Dutch societal perspective) in which we have conducted them, but they also contribute to the knowledge and insights from health economics.

Health economics and modelling

As a health economist, I often see the use of health economic models in order to guide decision-making, in many different contexts. Cost-effectiveness analyses in healthcare can bring great value, context, and guidance in the evaluation of treatments. They are often used in drug approval processes but offer value in many other decisions as well such as guideline development, analysing impact on society of new introductions, or long-term policy decisions. In times of budget and capacity constraints, we need data and research delivering most accurate information to take decisions on healthcare. Modelling aids in making these decisions by using (often multiple) data sources, combining them, and practically predicting the future based on all information currently available, providing information much quicker than waiting for all information to become available

in the real world. In order to check whether our models accurately reflect how the disease process runs in reality, it is entirely possible to compare model results to real-world observed results and patterns. This in turn is a learning mechanism for better simulation of disease processes and thus better modelling. Models provide options when real-world data is not always completely available or of high enough quality. Naturally, there are limitations and objections to the use of modelling in HE decision-making as well. We should be well-aware of the benefits, limitations, and ways to evaluate.

Modelling full treatment strategies

Simulating disease processes in models takes a very good understanding of said disease processes and the intricacies that come with it. The field of Rheumatology is the leading example in this thesis. Rheumatoid Arthritis (RA) and gout were the two diseases modelled in chapters 2, 3, and 4. Both are rheumatic diseases but with differing requirements for modelling. RA is a continuous disease with differing levels of disease activity, but always present. Gout can be characterized by long periods of no symptoms while the underlying disease mechanism is at work, until at one point a flare-up occurs and patients experience severe symptoms. Another important factor is the analysis of treatment strategies as opposed to single drugs, which requires accurate modelling and a vast source of data to properly feed a model with data points and transitions. For example, the Treat-to-Target (T2T) mechanism of treating rheumatoid arthritis, (chapters 2 and 3) is one of the key characteristics determining how the model is shaped. Another key consideration is whether to analyse a (combination of) treatments or a full treatment strategy or sequence.

Analysing treatment strategies as opposed to single drug (combinations), if applicable in the field of interest, can provide a more complete view of treatment of a certain disease process. In inflammatory chronic diseases, for example, this is of high importance. Patients with RA or gout will most likely require lifelong treatment. A certain treatment might work while others don't, this could also change over time and may differ between patients. The order in which those treatments are conducted matters greatly for long-term outcomes, in the case of RA, long-term damage is one of the risks of inadequate treatment in early stages. Modelling strategies naturally also requires more data inputs and modelling techniques. In RA, the subject of the HE analyses performed in chapters 2 and 3, treatment strategies and the context in which they are given are very important. Treatment sequences are important, as in lots of fields, more and more drugs or other treatment options are becoming available, and the order in which they are given can matter greatly for patient outcomes. For example, when guidelines do not (yet) distinguish between multiple treatment options within strategies, cost-effectiveness analyses can hugely benefit the choices to be made between different treatment strategies.

Aspects of Full HTA

In this dissertation, much attention has been paid to the cost & economic aspects of treatment strategies, chapters 2, 3, and 4 specifically. Assessment of these aspects can be conducted through cost-effectiveness analysis, budget impact analysis, burden of disease assessment, economic value assessment, and other options. Naturally, in cost-effectiveness and cost-utility analyses, effectiveness and an assessment of patient utility is considered as well. In this way, the trade-off

between costs and the added value of a treatment, treatment strategy, or other piece of health technology can be assessed. Integration of economic aspects, in combination with utility or effectiveness can lead to more informed decisions because of knowledge of cost and budget impact. This in turn could lead to more equitable distribution, and an optimization of resources in society.

Current use was addressed in chapter 6, within the context of a real-world setting pragmatic trial. The models in chapters 2, 3, and 4, were largely populated by real-world data, gathered in daily clinical practice in the Netherlands. Technical aspects were not a significant topic in this thesis. Safety aspects were not directly addressed in this thesis, naturally, this has been thoroughly researched in RCTs. This is similar to clinical aspects, these are covered by RCTs, chapter 6 also gives insight into clinical outcomes in a real-world context. Ethical considerations come into play when considering how the available healthcare budget should be allocated between healthcare services, disease areas, or treatment options. Cost-effectiveness analysis, like in chapters 2, 3, and 4, can aid in making these decisions. Decision-making and ethical considerations among the full spectrum of healthcare requires studies like these to be externally comparable. Organisational aspects are considered throughout this thesis to some extent due to the nature of treatment of RA and Gout, two chronic diseases, in which full strategies are evaluated. Legal aspects surrounding (the reimbursement of) expensive medications in, among others, rheumatology, have been a major topic of concern. At the introduction of TNFi, the financial aspect became a big topic due to high medication prices. Their introduction has triggered systems into place. This thesis does not research these specific mechanisms and legal and organisational aspects but they are important considerations throughout the other HTA framework dimensions as well.

Drug Prices in Economic Evaluations

Chapter 2 answered the question of whether a step-up therapy or an initial combination strategy in T2T-setting is cost-effective in the treatment of early RA patients. The results revealed that initial combination therapy is cost-effective and the dominant strategy. For these analyses, as health economists, we utilize list prices for treatments. Even though specific agreements are highly confidential, it is known that innovative drugs often come to market with negotiated deals on price with governments, pharmacy purchasing groups, individual hospitals, and insurers. What remains unknown, in the case of many cost-effectiveness studies, is how the results would change based on the actual prices that are being paid in the studied setting for medication. To have full and accurate impact with these analyses, ideally, we would be using actual pricing. Total budget and yearly cost of certain indications and treatments can only be realistically assessed if actual cost of all inputs would be available. Not only for medication, but for other aspects in the analysis as well. In chapter 2 and 3, we saw dominant strategies, meaning one strategy yielded higher utility at lower costs compared to the comparator strategy in that study. However, results could be influenced when taking into account real-world pricing. Chapter 3 of this thesis provides a possible option to give some insight into uncertainty around pricing in economic evaluations. In Chapter 2 displays a graph (i.e., Figure 3) in which the prices of the two main drugs (within a strategy) that were analysed can be varied. As a result, if one knows the actual price of the treatment strategies, one could then deduce the favourability of either strategy at said prices. While this graph provides

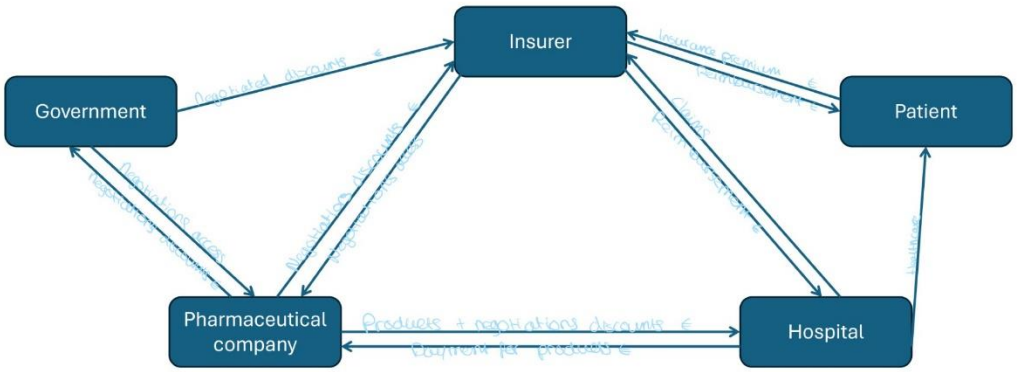
way of displaying a multitude of scenarios, it can only vary the price of two drugs in one graph. Graphs or other initiatives like these grant more information, to those who have the necessary info to interpret them. This increases transparency and could provide benefits if implemented more regularly. Additionally, providing more transparency on actual pricing openly would be even better, and would allow more parties and experts to make adequate conclusions.

(Lack of) Transparency in the Dutch system

If we use the Dutch setting as an example, there are many layers to the market in which hospitals, pharmaceutical companies, healthcare insurers and governments operate together. The schematic overview is sketched out in Figure 1 below. For drugs that fall into the hospital budgets in the base-case model, it would be a simple triangle. Hospitals (or hospital pharmacy buying groups, a collective buying organization) would buy medications from the pharmaceutical company. Once medication is given to a patient, hospitals would submit their declarations for reimbursement and the healthcare insurers would reimburse. Drugs that fall in the extramural category have a separate system. The vast majority of the treatments assessed in this dissertation fall under hospital budgets, the ones that do not have smaller impact, in terms of price. The negotiations between hospital and pharmaceutical company are subject to non-disclosure agreements (NDAs). Since 2015, the Dutch government is placing expensive treatments over certain thresholds in the so-called *lock.* After lock placement, procedures can be followed for advice in terms of cost-effectiveness, appropriate use and cost-reducing measures. This leads to careful consideration by multiple Dutch governmental organizations after which a financial arrangement is sought; the government enters negotiations with the pharmaceutical company bringing the product to market. Again, the result of these negotiations is confidential. However, the government does share average percentage of discounts negotiated over a full year. In 2019, the reported percentage across all price arrangements made was 42.4%. This has risen to 48.5% in 2022 [230]. Increased transparency could lead to lower prices and reduced inequity. The lack of transparency is due to several factors. Firstly, pharmaceutical companies have certain interests to protect, they could aim to keep prices confidential to prevent markets in different countries from influencing each other. Governments might want to maintain negotiating leverage and avoid administrative complexities. Both parties could be concerned about effects of negotiation transparency on various markets.

Another example of this is the negotiations between the collective of health insurers and pharmaceutical companies. For certain medications that do not cross the threshold for lock placement, the collective of Dutch healthcare insurers aims to come to a financial arrangement with the pharmaceutical company. The multiple layers and ways of making agreements on price make it a complex field. More transparency on price and process would enhance the applicability and accuracy of cost-effectiveness analyses, modelling, and broader health economic evaluations. International examples towards simplification should be considered.

Figure 1: Schematic overview of the Dutch intramural situation



Perspective – one size does not fit all

Chapter 4 of this thesis conducted analysis on cost-effectiveness of multiple gout treatments. The Dutch Guidelines for Economic Evaluation recommends using a societal perspective for cost-effectiveness analyses . This is advised as the societal perspective is considered to a rather complete perspective, as it takes into perspective many different factors [231]. This is the perspective that was chosen in chapters 2, 3, and 4, but the question remains whether this actually fits all.

The answer to which aspects in an economic evaluation are most important depends largely on who you ask. A hospital pharmacy will answer differently from a healthcare insurance company, there might even be differences when you ask two different department within a hospital. The ministry of Health, Welfare, and Sport will have a different perspective then the Ministry of Economic Affairs. A hospital may be more focused on capacity, in terms of staff and physical space. The hospital pharmacy might focus more on their budgets and thus the margin between reimbursed price for drugs and the price they have to pay towards the pharmaceutical industry. A health insurer, on the other hand will be more focused on management of claim costs and ensuring accessibility and affordability for all their insured individuals and might be less interested in work productivity loss. Whereas that point is of great importance to the Ministry of Economic Affairs.

The societal perspective, although comprehensive, may not always be sufficient. The Dutch healthcare system is divided in many different (financial) silos. In this context, the definition of one single “Societal Perspective” may be hard to define as different organizations, and even departments within those organization, are responsible for their own budgets, constraints, and incentives and thus will have differing perspectives. The Dutch societal perspective as a whole should be taken into considerations, but do to financial silos, where benefits might fall in a different budget than where the investment should be made, decisions are not taken with this comprehensive view in mind.

Free-market dynamics in the Dutch healthcare sector have led to the various stakeholders in the current system, each with their own interests. Each party is financed in their own way, which leads to their specific interests not necessarily being aligned with the interest of the broader society. This includes the ministry of healthcare, health insurers, pharmaceutical companies, hospitals, and

other healthcare organisations. In the field of rheumatology, the process of buying and reimbursement of expensive medications is complex and viewed differently by each stakeholder. The government and its subsidiary parties decided whether treatments, like TNFi, should be reimbursed. Healthcare insurers reimburse these drugs, while individual hospitals or buying groups negotiate with pharmaceutical companies on the exact (mostly confidential) prices.

If we do decide that decisions in healthcare policy-making should be taken with the societal perspective in mind, it does not make sense for these decision to be taken by parties that have financial and organisational incentives that do no align with the full scope of the societal perspective. Either these financial incentives need to change, to incorporate the societal perspective to a larger degree or the decision-making should be done by parties that do not have conflicts of interest in this scenario.

Data requirements

Analysing full treatment strategies also provides the opportunity to use a more comprehensive view of the impact of these decisions from a societal perspective. HE decisions can then take into account not only clinical outcomes, but also impact on overall costs, capacity, informal care, out-of-pocket expenses for patients, lost work productivity or wages, etc. Naturally, lots of data is needed to conduct this. Next to that, decisions have to be made, when carrying out HE analyses, what to analyse. This process requires lots of information and pragmatic choices. For example, in a case where we have 15 treatments for a certain condition, analysing every possible treatment (15! = 1,307,674,368,000) is, obviously, practically impossible. Additionally, data restraints are a huge limitation. Imagine not only having to code a trillion different strategies but also having enough adequate data to fill such a model.

A lot of those sequences will not make sense when presented to a medical doctor. We strongly rely on the vast amount of clinical knowledge, theoretical understanding of said treatments, patient characteristics, and patient history which all influence the number of feasible treatment sequences or strategies that are useful to evaluate. Evaluating the 'makes-sense' and relevant treatment strategies, however, brings much more perspective, context, and useful results than sticking to comparisons of single drugs. In order to do so, we do still rely heavily on availability of data. To give a complete picture and provide the aforementioned societal perspective of health economic impact of treatment (strategies), and demonstrated in this thesis, we need more than solely clinical measurements. In an ideal world, we would have access to perfect information and data on all these aspects that link to healthcare decisions.Possibly, this could be aided by artificial intelligence (AI) upcoming technologies. AI could integrate data from different sources, multiple electronic health records, apps, imaging devices, and more. One prerequisite is that data in these different systems contain variables which enable either human data managers or AI technologies to link the multiple data sources. Linked data sources can provide

Longitudinal patient-levels are absolutely essential to track disease progression and patterns. Especially in but not limited to, a field like rheumatology with chronic conditions. Longitudinal data allows healthcare professionals and researchers to assess effectiveness, costs, and safety of

treatment (strategies) over extended periods of time. To completely understand the health economic impact of different treatment strategies, we have to be able to link information on quality of life to outcomes, clinical outcomes, exact dosing, treatment (strategy), patient history, and costs. Accurate cost data is vital for meaningful analysis.

Clinical outcomes

Firstly, clinical outcomes, even though often recorded in hospital information systems, are not easily available by definition. In many cases, clinical outcomes data are being captured in free-text fields. This is a basic form of documentation and may work fine as long as the patient remains with the same physician, and this is the main reason for documenting. However, it is not very easy to put this way of data registration to other uses. Each physician may have an own way of exactly recording information. You can imagine how this would lead to differences not only between doctors in the same hospital but at a much larger scale it becomes increasingly difficult. More and more solutions are becoming available, among those are tools to *scrub* electronic patient dossiers and extract data from free-text fields are being used more and more. It is amazing to see how technology can provide solutions. However, they do remain suboptimal as compared to structured registration at the source. It can be difficult to connect information from these free-text fields to other information from a different system within the hospital like current treatment and exact dosing. Generally, it would be easier if outcomes like these are registered, at the source, in a way that is very straightforward to utilize in daily clinical practice and standardized to connect to other data points, and to analyse by researchers.

Costs

Secondly, information on costs incurred by the patient is needed. This relates to medication cost, healthcare costs both directly related and unrelated to the disease of interest (including GP or specialist visits, but also imaging or other procedures), and out-of-pocket (OOP) costs. Another category of costs that is of great importance is the loss of work productivity, or sick days incurred. These last two categories, OOP costs and work productivity loss are two inputs that, when collected, have to come directly from the patient themselves, whereas medication costs and other healthcare costs can be derived from patient records. The research in this dissertation was able to extract extensive information on costs from the DREAM registries.

HRQOL

Thirdly, health-related quality of life (HRQOL) should be recorded. This can be done with different instruments all have benefits and drawbacks. The specific instrument used to capture HRQOL should be considered carefully and is highly relevant as it leads to input for decision-making in by health policy makers and possibly physicians as well. HRQOL can be captured in various ways, questionnaires can differ a lot in length, the EuroQol five-dimension (EQ-5D) consists of 5 questions, whereas a Short-Form 36-item health survey (SF-36) consists of many more questions. This all leads to a different burden for patients. When choosing which questionnaire or method to use, researchers and medical professionals should be very aware of burden on patients and the health system as well as methodological differences. It is important to consider burden on patients and external comparability between multiple studies when choosing between different

instruments capturing HRQOL. It should be noted that introductions of new instruments should be carefully considered as well. As outlined in Chapter 5, responsiveness of these questionnaires can differ as well. It should be noted that responsiveness to a decline in health state is not automatically equivalent to an instrument's responsiveness to improvements in a patient's health state. Commonly-used HRQOL instruments lack responsiveness in case of deterioration in health state. What is needed is either improved instruments or a better understanding of and caution when using these instruments, specifically in certain cases. When health economists and policymakers need to take decisions on disinvestment, revaluation of already approved treatments, or willingness-to-accept (WTA) scenarios, we should be very aware of the limitations that some HRQOL instruments have in picking up worsening of patient's health state. This could lead to an underestimation of the degree to which the patient has actually worsened and lead to decision-making with suboptimal information. In cost-effectiveness analyses, this could lead to an incorrect assessment of patient HrQOL and thus incorrect quality-adjusted life years (QALY), leading to results that might not fully reflect reality. If we are aware of these limitations and considerations though, they can be taken into account. This phenomenon may be a result of such advancements in healthcare to a point where, in certain medical fields, healthcare providers are less focused on solely addressing rather severe health states but are finetuning or optimizing treatment when patients are in relatively better condition. The introduction of innovative treatment (strategies) in rheumatology has certainly contributed to this shift. Accurate responsiveness to deterioration in health state is not the same as responsiveness to improvements and we should be aware of that. As an analogy, a thermometer that measures freezing temperatures is not designed in the same way as one designed to measure to caramelization point of sugar (at around 170°C). If we do want to accurately measure responsiveness in the full spectrum of health states we have to design improved instruments.

Measuring health or disease?

When collecting patient-level data, and more specifically (clinical) outcomes and HRQOL, in the field of rheumatology it is important to distinguish between measuring health versus measuring disease. Both approaches serve different purposes and is utilized by stakeholders, like physicians, health economists, or policy makers, in different ways. The differences between these two approaches have significant implications on patient experience, design of study or registry, selection of instruments, and the use of data.

Measuring health relates to a comprehensive way of assessing a patient's status. It emphasizes mental and physical well-being. Social relationships, ability to execute daily tasks, physical functioning can all be a part of this. Health, in this sense, is measured by generic instruments like the EQ-5D or SF-36. These tools focus on multiple dimensions that combined make up a full picture of how a patient is doing. Results have a broad applicability and allows for comparison across multiple disease areas and populations. For policy makers, this allows to compare results from different medical fields and see a comparable overview of impact on patients of various treatments or interventions. It also helps to quantify and create an overview of the burden of disease in different indications, leading to helpful information in long-term healthcare investment decisions.

Measuring disease would relate more to the measurement of pain, fatigue, joint function or swelling (for example, in rheumatology), lung function (for example, in asthma) or status of disease. The applicability is narrower and more specific to a particular disease or disease area. It is more often utilized by specialized physicians and nurses.

Both approaches have benefits and drawbacks but mostly, they each have their own role and are crucial in health economic analyses. For example, in chapters 2-4, we have utilized both perspectives to allow for the most accurate models to simulate disease processes. A balanced approach allows for all perspectives to be taken into account. At the same time, the administrative burden on patients as well as healthcare professionals must be considered when decisions for (multiple) questionnaires to be filled in are taken.

Administrative burden or joy

For all types of data—whether clinical outcomes, cost information, or utility measures—accurate and timely registration is absolutely essential. If this data is not properly recorded, it becomes unusable, which represents a significant missed opportunity for meaningful analysis and improvements in patient care. But moreover, it represents a huge loss in terms of time and effort that was put into recording the information. We have to make sure that when we put time, resources, and effort into data collection (by patients and healthcare professionals) we can actually utilize these data to improve healthcare, and not waste their efforts. The administrative burden for health professionals and patients must remain as low as possible, while still providing the right data to improve quality of care.

I argue that with well-thought out systems and proper integration, we can generate better data while reducing the administrative burden. The technological possibilities are enormous. Processes that can be explained can be coded. A frequently mentioned term is ‘privacy by design’. One could argue that we also need to prioritize the continued usability of data by design. Every added step or action that we place upon healthcare professionals and patients should be well-considered and serve a specific purpose, particularly in times of budget and healthcare capacity constraints. The last thing that should be done is adding administrative burden to the plates of healthcare professionals when its usefulness could be debated. By properly designing these systems, privacy and safety is ensured, the system can provide the first checks on data quality, and as much as possible can be automated.

One possible avenue and set of guidance towards this is the concept of FAIR data, FAIR stands for Findable, Accessible, Interoperable, and Reusable data. FAIR is a framework ensuring that data is adequately managed and shared so that it maximizes usability, reusability and interoperability across fields, disciplines, and platforms. The concept of good data stewardship and what is necessary from that point of view has naturally been discussed and implemented for years. The FAIR concept can be viewed as a way of providing guidance and transparency in how to execute this and make sure that data can be used for many different purposes. These guidelines or best practices are available to anyone wanting to commit to the concept. It can provide transparency in scientific research, enhance visibility of research output, promote collaboration internationally and between research groups, and accelerate innovation.

Real-World Data Supplements Randomised Controlled Trials

In the fields of clinical research and health economics, the two primary sources of information on drug effectiveness are either randomised controlled trials (RCTs) or Real-World Data (RWD). RCTs are viewed as the gold standard when it comes to generating evidence on treatment efficacy. RCTs are conducted in a highly controlled manner and are designed to establish causal relationships between treatment and outcome. A meticulous protocol is followed, with strict in- and exclusion criteria, randomization to either intervention or control group, and everything is done to minimize any possible bias. In this way, confounding factors are minimised and we can best isolate the true effect of the studied drug or treatment. Yet, this approach also brings forward limitations. Certain population groups are often kept out of scope in RCTs, including elderly patients, children or teenagers, people with comorbidities, people that could get pregnant or are breastfeeding, among others. This last category deserves some additional attention. Women who are pregnant are often excluded from trials, due to label restrictions. These exclusion criteria can be expanded to women who are wanting to get pregnant or even further to anyone who could get pregnant and is not using double anti-conception. In this way, gender discrimination is perpetuated in healthcare by limiting access to clinical trials. The systematic exclusion of patients based on reproductive factors contributes to broader problem of gender bias and diversity in the broadest sense in healthcare.

The goal is to isolate the studied drug effect. However, these patients will at some point be within the intended treatment group once a drug is brought to market. High internal validity is one of the most important benefits of RCTs, while external validity or generalizability might lack. In order to truly know what innovations do and how they can best help all people that need them, we also need information on effectiveness and outcomes in the real world, including all complexities that arise there.

Real world evidence (RWE) is derived from data collected outside of these controlled trial settings, in daily clinical practice. This can include data from registries, electronic patient dossiers, insurance claims, or observational studies. RWE encompasses the complexity of all types of patients with vastly diverse backgrounds and characteristics but in a typical healthcare setting. These patients may have comorbidities and/or take multiple drugs. These factors provide complexities in registration, analysis and interpretation due to the less-controlled nature in which data is captured. Comparative analysis, selection bias, and data interpretation need to be carefully considered when using RWD. Chapters 2, 3, and 4 of this thesis rely heavily on the use of RWD, through which we were able to simulate disease processes in gout and RA, largely based on data collected in daily clinical practice.

Both RCTs and RWE have important roles in health economic evaluations and clinical decision-making. RCTs remain the gold standard and can establish efficacy of new treatments under controlled conditions. RWE can add to that by providing insights into effectiveness, safety, and health economic impact of treatments in daily clinical practice and within a broader context. Chapter 6 of this thesis generated real-world evidence.

Pragmatic trials like these naturally embody some limitations, but mostly provide an extra tool to gather more information on treatments in a fuller context and can provide value to patients,

healthcare providers, and health policy makers. Furthermore, RWE can provide valuable insight into the balance between costs and added value in terms of health. In the real world treatment of RA, the group of patients using the drug are more diverse, the duration of use is often longer, there are more comorbidities and also other drugs are used as well. This is the reality in which the studied drug is given, so this is the context in which we must find out how the treatment works for patients. Value-based real-world evidence can thus provide much more insight and applicability to the subsequent real world than RCT data. Chapter 6 has shown that real-world data and evidence is different from what is seen in pivotal trials. Comparative effectiveness in this setting is not necessarily the exact same as what was observed in RCTs. RWE is a valuable addition to RCT evidence and reflects the settings in which all types of patients are treated. This links strongly to point of cost-effectiveness studies and from which perspective they should be analysed. It can be argued that it is imperative that cost-effectiveness studies with a societal perspective should utilize data from a real-world setting, in order to provide results that translate to the societal perspective and real-world setting in which the results are to be used.

Imagine the possibilities of improved data availability

In order to conduct robust health economic evaluations, we need data on costs, utility, and clinical outcomes from real-world settings. Data must be longitudinal, correctly linked to treatments, strategies, comorbidities, and patient history in order to capture the true impact of healthcare interventions from different perspectives. This would be of great benefit for patients, healthcare professionals, payers, and decision-makers. Cost-effectiveness studies could be analysed from a societal real-world perspective, budget impact studies could be much more applicable to the real world and modelling becomes more accurate. The cost-effectiveness studies in this thesis are an illustration. A lot less assumptions are necessary, for example in terms of treatment effectiveness in patients that have been treated previously with other drugs. This applies not only in the field of rheumatology, but much broader as well. Unfortunately, initiatives building towards this often start small, and rely heavily on determination and motivation of healthcare professionals and researchers, funded by (temporary) grants or research awards. This leads to a certain momentariness and difficulty in upscaling projects, which is a shame. One could argue that all medical fields would benefit from a combined initiative like these, and it would have to be a(n) (inter)national endeavour.

The good news is, this dream scenario is entirely feasible.

While the project to achieve this may seem daunting, it is far from impossible. In fact, the feasibility has been demonstrated in, among other projects, the Dutch Rheumatoid Arthritis Monitoring (DREAM) registry. DREAM successfully collected real-world data, linked across multiple indications and treatment strategies, registered meaningful outcomes and utilized this information for impactful analyses. This demonstrates that when systems are properly integrated, collecting RWD at this detailed level is entirely possible and achievable. Combined with methodologically robust analysis, impact can be great.

The advantages of implementing systems like these are grand. Patients benefit from informed treatment decisions and access to their own health status information. Healthcare professionals

are able to have more insight in their own patterns of treatment, possibly improving their quality of care. For society, the ability to assess health outcomes leads to improved resource allocation and potentially lowers healthcare costs. Policy makers have access to reliable evidence that inform decision making. Additionally, it would allow for benchmarking of healthcare providers on their quality of care. This would allow physicians to learn and improve and allow patients and policymakers to assess quality and make decisions based on that.

Transparency in outcomes

It is feasible at reasonable costs. A national initiative like this does require investment, but it is financially feasible. Contrary to popular belief, implementing large-scale data collection systems does not necessarily mean exorbitant costs. In the past, collecting data was rather time- and cost-consuming. However, by leveraging existing digital infrastructure, such as electronic patient dossiers, and making thoughtful investments in interoperable systems, it is possible to keep costs under control. The level of automation has increased and is still increasing. Additionally, there is a level of duplication and redundancy that could be eradicated. It should be noted that validity checks by clinical experts remain necessary.

If we now take a more holistic view of the health economic landscape, we can only imagine a society where a concept like this is the blueprint. This would be a dream healthcare data scenario with added value for all involved. Data-driven healthcare for all should be the norm, improving quality of care and keeping affordability and accessibility in high regard.

Acknowledgements

After the general discussion, I want to make sure to give the proper attention and appreciation to the people who have made the collection of all data used in this thesis possible.

These are all the patients that were included in DREAM, pivotal trials, health economic evaluation trials, and all who have been subject to questionnaires or any other form of data collection. Without any of you, none of this would have been possible, thank you. Thank you for answering those questions, thank you for filling out questionnaires, thank you for agreeing to be part of a study or observational cohort, thank you for contributing to research.

To the rheumatologists, physicians, nurses, supportive staff in hospitals and other health care organisations, thank you for all your efforts in healthcare, thank you for taking care of patients and improving their quality of life, thank you for collecting data, thank you for all that you do.

We are not only standing on the shoulders of giants in terms of prior researchers (thank you!) but we are also truly standing on the shoulders of everyone working in healthcare, on whose enthusiasm and relentless motivation towards improving quality of care we all rely.

My thanks are also due in a large part to all the academic colleagues I have had the privilege of working with other the past years. Thank you for your contributions to this research. You have inspired, challenged, and motivated me. A special thank you to my promotors and committee members, thank you for your invaluable contributions to this thesis.

Summary

Rheumatoid arthritis (RA) and gout are chronic diseases. Both face their own sets of challenges in diagnosis and treatment. Advancements in therapies have made management easier and effective. These therapies have improved patient outcomes but also driven costs higher than before. Health economic modelling has helped to evaluate cost-effectiveness of these treatments. Utilizing real-world data (RWD) paired with modelling to simulate disease processes can increase our understanding of treatment outcomes, disease progression, and societal impact and can guide healthcare policy with data-driven decisions. One very important and debated aspect of RWD is its quality and analytical rigor. Instrument responsiveness, (technical) database management, and proper methodology are of vital essence.

Chapter 2 compared two treatment strategies for rheumatoid arthritis, both Treat-to-Target strategies by adapting and improving upon a previously validated Markov Model. Ultimate freedom for simulated patients was achieved in this model, in concordance with patterns observed in daily clinical practice, including switching between all states and tapering medication. Step-up therapy was compared to initial combination therapy using a markov model. Input was retrieved from Dutch Rheumatology Monitoring registry (DREAM) Remission Induction Cohort I (step-up) and cohort II (initial combination). Expected costs and utility after five years were compared between the two strategies. Initial combination therapy was found to be cost-effective and the dominant strategy. This was explained by higher remission percentages in the initial combination strategy at every time point, yielding more utility at a lower cost in treating recently diagnosed RA patients to the target of remission.

Chapter 3 evaluated the cost-effectiveness of two treatment strategies for RA again in a Treat-to-Target set-up. The European League Against Rheumatism (EULAR) and American College of Rheumatology (ACR) guidelines were inconclusive about the distinction between Janus Kinase inhibitor (JAKi) versus biological disease-modifying anti rheumatic drug (bDMARDs) after failure to reach remission in the first step in treatment strategy. The first treatment had been implemented as standard of care and RWD was available. The second treatment strategy, involving JAKi, had not yet generated such data. In order to assess Real-World cost-effectiveness, RWD and data from Randomised Controlled Trials (RCTs) was combined. Naturally, to do so in a methodologically sound manner, specific methods had to be applied. The relative risk method was used to adequately combine data sources, adjusting RCT data to be correct in use next to RWD sources in this model. A Markov model was constructed to compare the two strategies utilizing data from DREAM and pivotal trial data, using Monte Carlo simulation and probabilistic sensitivity analysis. The JAKi strategy was shown to be cost-effective in conventional synthetic disease-modifying anti rheumatic drug (csDMARDs) refractory patients in a Monte Carlo simulation.

Chapter 4 assessed the cost-effectiveness of different treatments in gout. Different combinations of urate lowering treatment (ULT) and inflammatory treatments were analysed using a Markov model. Health states in this model were defined by pain level, which was a novel approach. Gout is a disease where long periods of no or few symptoms are experienced, even though the underlying disease mechanism is continuing. These so called intercritical periods alternate with

periods of flares, in which extreme pain is experienced. This intermittent disease process was simulated in a Markov model. Different combinations of treatment, including the option to not use ULT, were considered. Model input was taken from patient-level clinical trial data, meta-analyses, and previous health-economic evaluations. Combinations using allopurinol and febuxostat as ULT were cost-effective. Combinations not using any ULT were not cost-effective. With respect to inflammatory treatment, combinations using colchicine, naproxen and prednisone all offered comparable health economic implications, naproxen was the ultimate favoured option.

Chapter 5 evaluated the responsiveness of generic and disease-specific instruments in patients experiencing a decline in health status. The analyses were conducted in a dataset of RA patients that were in stable remission before discontinuing their treatment. This was done in a 12-month trial. Tools like the disease activity score in 28 joints (DAS28), EuroQol-Five Dimension (EQ-5D), and Short-Form Six Dimension (SF-6D) were assessed, among others. Standardized response means (SRM), receiver operating characteristic (ROC) area under the curve (ROC-AUC), and relative validity (RV) were used to assess responsiveness in these instruments. The analyses revealed that the instruments studied did detect changes in patients' health state, however, sensitivity falls short. There is a need for improved understanding of adequately capturing changes in health status in case of deterioration.

Chapter 6 compares the use of baricitinib, a JAKi with use in tumor necrosis factor alpha inhibitors (TNFi) in patients with RA, who had not responded to first line of treatment, csDMARDs in a Treat-to-Target context. This was an open-label, pragmatic, real-world setting trial, named PERFECTRA. 199 patients were included randomized to receive baricitinib or a TNFi. Baricitinib displayed non-inferiority and subsequent superiority at the primary endpoint: achieving American College of Rheumatology 50 (ACR5) response at 12 weeks (42% vs. 20%). Secondary endpoints 28-joint count Disease Activity Score with C reactive protein (DAS28-CRP)<2.6, changes in PROMs and radiographic progression. Baricitinib strategy also demonstrated comparable or better outcomes in those compared to TNFi strategy. This study suggested baricitinib could be an improved option as compared to TNFi in patients with RA who did not respond to csDMARDs.

In all healthcare systems, decisions must be made continuously. It needs not to be discussed that healthcare suffers from complexity. Management information is generally good and sometimes abundantly available, but usable outcome information remains a major challenge. Improved data recording and infrastructure can allow healthcare professionals, researchers, and policymakers to make decisions based on (more) accurate information and conclusions and cost-effectiveness modelling accuracy can improve. Cost-effectiveness analyses can lead to improved decision-making in healthcare. For adequate analysis from a societal perspective, adequate, accurate and a vast amount of data is needed.

Samenvatting

Reumatoïde artritis (RA) en jicht zijn chronische ziekten die elk hun eigen uitdagingen kennen op het gebied van diagnose en behandeling. Ontwikkeling in therapieën heeft het behandelen van deze aandoeningen makkelijker en effectiever gemaakt. Deze therapieën hebben de patiëntuitkomsten verbeterd maar hebben ook geleid tot verhoogde kosten. Gezondheidseconomisch modelleren heeft geholpen om de kosteneffectiviteit van deze behandelingen en behandelstrategieën te evalueren. Het gebruik van real-world data (RWD) in combinatie met modellering om ziekteprocessen te simuleren kan ons begrip van behandeluitkomsten, ziekteprogressie en maatschappelijke impact vergroten en kan het gezondheidsbeleid sturen met data-gestuurde beslissingen. Een belangrijk en veel besproken aspect van RWD is de kwaliteit en analytische nauwkeurigheid. De responsiviteit van instrumenten, (technisch) databasemanagement en een correcte methodologie zijn van essentieel belang.

Hoofdstuk 2 vergeleek twee behandelstrategieën voor reumatoïde artritis, beide Treat-to-Target-strategieën, door een eerder gevalideerd Markov-model aan te passen en te verbeteren. In dit model werd grote vrijheid voor gesimuleerde patiënten bereikt, in overeenstemming met patronen die worden gezien in de dagelijkse klinische praktijk, waaronder het bewegen tussen alle *health states* en ook het afbouwen van medicatie. Step-Up therapie werd vergeleken met initiële combinatietherapie. De input werd verkregen uit de Nederlandse Reumatologie Monitoring (DREAM) Remission Induction Cohort I (step-up) en Cohort II (initiële combinatie). Verwachte kosten en utiliteit na vijf jaar werden tussen de twee strategieën vergeleken. Initiële combinatietherapie bleek kosteneffectief en de dominante strategie te zijn. Dit werd verklaard door hogere remissiepercentages bij de initiële combinatietherapie op elk moment, wat meer utiliteit opleverde tegen lagere kosten bij de behandeling van recent gediagnosticeerde RA-patiënten met als doel remissie.

Hoofdstuk 3 onderzocht de kosteneffectiviteit van twee behandelstrategieën voor RA binnen een Treat-to-Target-aanpak. De richtlijnen van de European League Against Rheumatism (EULAR) en het American College of Rheumatology (ACR) waren niet eenduidig over het onderscheid tussen Janus Kinase-remmers (JAKi) en biological disease-modifying anti rheumatic drug (bDMARDs) na falen in het bereiken van remissie in de eerste stap van de behandelstrategie. De eerste behandeling was geïmplementeerd als standaardzorg en RWD was beschikbaar. Voor de tweede behandelstrategie, waarbij JAKi werd gebruikt, waren dergelijke gegevens nog niet beschikbaar. Om de kosteneffectiviteit in de realiteit te beoordelen, werden RWD en gegevens uit gerandomiseerde gecontroleerde onderzoeken (RCT's) gecombineerd. Natuurlijk moesten hiervoor specifieke methoden worden toegepast om methodologisch correct te blijven. De relatieve risico-methode werd gebruikt om gegevensbronnen adequaat te combineren, waarbij RCT-gegevens werden aangepast voor correct gebruik naast RWD in dit model. Een Markov-model werd geprogrammeerd om de twee strategieën te vergelijken, met gebruik van gegevens uit DREAM en gegevens uit randomised controlled trials via Monte Carlo-simulatie en probabilistische sensitiviteitsanalyse (PSA). De JAKi-strategie bleek kosteneffectief bij patiënten die niet reageerden

op conventionele synthetische ziekte-modificerende anti-reumatische geneesmiddelen (csDMARDs) in een Monte Carlo-simulatie.

Hoofdstuk 4 beoordeelde de kosteneffectiviteit van verschillende behandelingen voor jicht. Verschillende combinaties van urinezuurverlagende behandelingen (ULT) en ontstekingsremmende behandelingen werden geanalyseerd met behulp van een Markov-model. Gezondheidstoestanden in dit model werden gedefinieerd op basis van het pijnniveau, wat een nieuwe benadering was. Jicht is een ziekte waarbij lange perioden van weinig of geen symptomen worden ervaren, terwijl het onderliggende ziektemechanisme doorgaat. Deze zogenaamde interkritische perioden wisselen af met perioden van opvlammingen, waarin extreme pijn wordt ervaren. Dit intermitterende ziekteproces werd gesimuleerd in een Markov-model. Verschillende behandelcombinaties, inclusief de optie om geen ULT te gebruiken, werden overwogen. De input van het model werd verkregen uit patiëntgegevens op proefniveau, meta-analyses en eerdere gezondheidseconomische evaluaties. Combinaties met allopurinol en febuxostat als ULT bleken kosteneffectief te zijn. Combinaties zonder ULT waren niet kosteneffectief. Met betrekking tot ontstekingsremmende behandeling boden combinaties met colchicine, naproxen en prednison vergelijkbare gezondheidseconomische implicaties, waarbij naproxen de uiteindelijke favoriete optie was.

Hoofdstuk 5 evalueerde de responsiviteit van generieke en ziekte-specifieke instrumenten bij patiënten die een verslechtering van hun gezondheidstoestand ervoeren. De analyses werden uitgevoerd in een dataset van RA-patiënten die in stabiele remissie verkeerden voordat zij hun behandeling stopzetten. Dit werd gedaan in een 12 maanden durende trial. Instrumenten zoals de Disease Activity Score in 28 gewrichten (DAS28), EuroQol-Five Dimension (EQ-5D) en Short-Form Six Dimension (SF-6D) werden geëvalueerd, naast andere. Gestandaardiseerde responsmiddelen (SRM), receiver operating characteristic (ROC) area under the curve (ROC-AUC) en relatieve validiteit (RV) werden gebruikt om de responsiviteit van deze instrumenten te beoordelen. De analyses toonden aan dat de onderzochte instrumenten veranderingen in de gezondheidstoestand van patiënten detecteerden, maar dat de gevoeligheid tekortschiet. Er is behoefte aan een verbeterd begrip van hoe veranderingen in de gezondheidstoestand bij verslechtering adequaat kunnen worden vastgelegd.

Hoofdstuk 6 vergeleek het gebruik van baricitinib, een JAKi, met tumor necrose factor alfa-remmers (TNFi) bij patiënten met RA die niet hadden gereageerd op de eerstelijnsbehandeling, csDMARDs, binnen een Treat-to-Target-context. Dit betrof een open-label, pragmatische, real-world setting trial genaamd PERFECTRA. In totaal werden 199 patiënten gerandomiseerd om baricitinib of een TNFi te ontvangen. Baricitinib toonde non-inferioriteit en daaropvolgende superioriteit aan op het primaire eindpunt: het behalen van de American College of Rheumatology 50 (ACR50) respons na 12 weken (42% versus 20%). Secundaire eindpunten omvatten een DAS28-CRP-score <2,6, veranderingen in PROM's en radiografische progressie. De baricitinib-strategie demonstreerde ook vergelijkbare of betere uitkomsten in vergelijking met de TNFi-strategie. Deze studie suggereert dat baricitinib een verbeterde optie kan zijn vergeleken met TNFi bij RA-patiënten die niet reageren op csDMARDs.

In alle gezondheidszorgsystemen moeten voortdurend beslissingen worden genomen. Het is duidelijk dat de gezondheidszorg complexiteit kent. Managementinformatie is over het algemeen goed en soms overvloedig beschikbaar, maar bruikbare uitkomstinformatie blijft een grote uitdaging. Verbeterde gegevensregistratie en infrastructuur kunnen zorgverleners, onderzoekers en beleidsmakers in staat stellen beslissingen te nemen op basis van (meer) nauwkeurige informatie en conclusies, en de nauwkeurigheid van kosteneffectiviteitsmodellen kan worden verbeterd. Kosteneffectiviteitsanalyses kunnen leiden tot betere besluitvorming in de gezondheidszorg. Voor een adequate analyse vanuit maatschappelijk perspectief zijn goede, nauwkeurige en uitgebreide gegevens nodig.

References

1. Allen A, Carville S, McKenna F. Diagnosis and management of rheumatoid arthritis in adults: Summary of updated NICE guidance. BMJ (Online) 2018;362:1–4.

2. Venetsanopoulou AI, Alamanos Y, Voulgari P V., Drosos AA. Epidemiology and Risk Factors for Rheumatoid Arthritis Development. Mediterr J Rheumatol2023;34:404–13.

3. Belbasis L, Dosis V, Evangelou E. Elucidating the environmental risk factors for rheumatic diseases: An umbrella review of meta-analyses. Int J Rheum Dis 2018;21:1514–24.

4. Fries JF, Hochberg MC, Medsger Jr TA, Hunder GG, Bombardier C, COMMITTEE ACODANDT. Criteria for rheumatic disease. Different types and different functions. Arthritis Rheum 1994;37:454–62.

5. Mrid R Ben, Bouchmaa N, Ainani H, El Fatimy R, Malka G, Mazini L. Anti-rheumatoid drugs advancements: New insights into the molecular treatment of rheumatoid arthritis. Biomedicine & Pharmacotherapy 2022;151:113126.

6. Smolen JS, Aletaha D, Koeller M, Weisman MH, Emery P. New therapies for treatment of rheumatoid arthritis. The lancet 2007;370:1861–74.

7. Prasad P, Verma S, Surbhi, Ganguly NK, Chaturvedi V, Mittal SA. Rheumatoid arthritis: advances in treatment strategies. Mol Cell Biochem 2023;478:69–88.

8. Lindqvist E, Saxne T, Geborek P, Eberhardt K. Ten year outcome in a cohort of patients with early rheumatoid arthritis: health status, disease process, and damage. Ann Rheum Dis 2002;61:1055–9.

9. Boini S, Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. Ann Rheum Dis 2001;60:817–27.

10. NVR, NHG. Richtlijn Diagnostiek van Reumatoïde Artritis.

11. Van METO. Conceptrichtlijn Medicamenteuze Behandeling van Reumatoïde Artritis.

12. Bijlsma Johannes WJ, Eric H. EULAR Textbook on Rheumatic Diseases. Italy. BMJ 2018;594–622.

13. van de Laar M, Hoentjen F, Bing Thio H. Nieuwe stap in behandeling van immuunziekten: Jakinibs, remmers van het intracellulaire Janus-kinase. Ned Tijdschr Geneeskd 2019;163.

14. Vermeer M, Kievit W, Kuper HH, Braakman-Jansen LMA, Moens HJB, Zijlstra TR, et al. Treating to the target of remission in early rheumatoid arthritis is cost-effective: results of the DREAM registry. BMC Musculoskelet Disord 2013;14:1–9.

15. 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. Ann Rheum Dis 2014;73:492–509.

16. Dalbeth N, Merriman TR, Stamp LK. Gout. The Lancet 2016;388:2039–52.

17. Singh JA. Challenges faced by patients in gout treatment: a qualitative study. J Clin Rheumatol 2014;20:172.

18. Kuo CF, Grainge MJ, Zhang W, Doherty M. Global epidemiology of gout: prevalence, incidence and risk factors. Nat Rev Rheumatol 2015;11:649–62.

19. Ragab G, Elshahaly M, Bardin T. Gout: An old disease in new perspective–A review. J Adv Res 2017;8:495–511.

20. Cronstein BN, Sunkureddi P. Mechanistic aspects of inflammation and clinical management of inflammation in acute gouty arthritis. Journal of Clinical Rheumatology2013;19:19–29.

21. Gonzalez EB. An update on the pathology and clinical management of gouty arthritis. Clin Rheumatol 2012;31:13–21.

22. Malik A, Schumacher HR, Dinnella JE, Clayburne GM. Clinical diagnostic criteria for gout: comparsion with the gold standard of synovial fluid crystal analysis. JCR: Journal of Clinical Rheumatology 2009;15:22–4.

23. Yeap SS, Gun SC. Gout: why is this curable disease so seldom cured? Ann Rheum Dis 2013;72:e5–e5.

24. Janssen CA, Oude Voshaar MAH, Vonkeman HE, Jansen TLTA, Janssen M, Kok MR, et al. Anakinra for the treatment of acute gout flares: a randomized, double-blind, placebo-controlled, active-comparator, non-inferiority trial. Rheumatology 2019;

25. Richette P, Doherty M, Pascual E, Barskova V, Becce F, Castañeda-Sanabria J, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis2017;

26. Kiltz U, Smolen J, Bardin T, Solal AC, Dalbeth N, Doherty M, et al. Treat-to-target (T2T) recommendations for gout. Ann Rheum Dis 2017;76:632–8.

27. Valkenburg HA, Ball J, Burch TA, Bennett PH, Lawrence JS. Rheumatoid factors in a rural population. Ann Rheum Dis 1966;25:497.

28. Van Leeuwen MA, Van Rijswijk MH, Sluiter WJ, Van Riel P, Kuper HH, Van de Putte LBA, et al. Individual relationship between progression of radiological damage and the acute phase response in early rheumatoid arthritis: towards development of a decision support system. 1997;

29. Tak PP, Taylor PC, Breedveld FC, Smeets TJM, Daha MR, Kluin PM, et al. Decrease in cellularity and expression of adhesion molecules by anti–tumor necrosis factor α monoclonal antibody treatment in patients with rheumatoid arthritis. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology 1996;39:1077–81.

30. ten Klooster PM, Veehof MM, Taal E, van Riel PLCM, van de Laar MAFJ. Changes in priorities for improvement in patients with rheumatoid arthritis during 1 year of anti-tumour necrosis factor treatment. Ann Rheum Dis 2007;66:1485–90.

31. Ministerie van Vorlksgezondheid W en S (VWS). Volksgezondheid en Zorg - Reumatoïde artritis (RA) Zorg Cijfers.

32. NZa. Kerncijfers intramurale dure geneesmiddelen [Internet]. 2024 [cited 2024 Nov 13]. Available from: https://www.nza.nl/zorgsectoren/geneesmiddelenzorg/kerncijfers-geneesmiddelenzorg

33. Eysink PED, Poos M, Gijsen R, Kommer GJ, van Gool CH. Epidemiologische data van ziekten van het botspierstelsel en bindweefsel: achtergrondrapport voor programma zinnige zorg. 2019;

34. Papanicolas I, Woskie LR, Jha AK. Health care spending in the United States and other high-income countries. JAMA 2018;319:1024–39.

35. World Health Organization. World Health Organization Global Health Expenditure Database. https://apps.who.int/nha/database2024;

36. Sorenson C, Drummond M, Bhuiyan Khan B. Medical technology as a key driver of rising health expenditure: disentangling the relationship. ClinicoEconomics and outcomes research 2013;223–34.

37. Lapointe L, Mignerat M, Vedel I. The IT productivity paradox in health: A stakeholder's perspective. Int J Med Inform 2011;80:102–15.

38. Porter ME, Teisberg EO. Redefining health care: creating value-based competition on results. Harvard business press; 2006.

39. Kristensen FB, Lampe K, Wild C, Cerbo M, Goettsch W, Becla L. The HTA Core Model®—10 years of developing an international framework to share multidimensional value assessment. Value in Health 2017;20:244–50.

40. Lampe K, Mäkelä M, Garrido MV, Anttila H, Autti-Rämö I, Hicks NJ, et al. The HTA core model: a novel method for producing and reporting health technology assessments. Int J Technol Assess Health Care 2009;25:9–20.

41. Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in ecomomic evaluation: an unavoidable fact of life. Health Econ 1997;6:217–27.

42. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics 1998;13:397–409.

43. Hatswell AJ, Bullement A, Briggs A, Paulden M, Stevenson MD. Probabilistic sensitivity analysis in cost-effectiveness models: determining model convergence in cohort models. Pharmacoeconomics 2018;36:1421–6.

44. Thom H. Deterministic and probabilistic analysis of a simple Markov model: how different could they be? Appl Health Econ Health Policy 2022;20:447–9.

45. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. Medical decision making 1985;5:157–77.

46. Van Hout BA, Al MJ, Gordon GS, Rutten FFH. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 1994;3:309–19.

47. Al MJ. Cost-effectiveness acceptability curves revisited. Pharmacoeconomics 2013;31:93–100.

48. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ 2001;10:779–87.

49. O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. Med Care 1994;150–63.

50. The_EuroQol_Group. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy (New York) 1990;16:199-208.

51. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.

52. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459-68.

53. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PMM. On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument evaluation. Quality of Life Research 2003;12:349–62.

54. Sweeney S, Harris E, Firestein G, Jeffery RC, Venables PJW MR. Clinical features of rheumatoid arthritis. In: Kelley's Textbook of Rheumatology. 2013. page 1109–36.

55. Huscher D, Mittendorf T, Von Hinüber U, Kötter I, Hoese G, Pfäfflin A, et al. Evolution of cost structures in rheumatoid arthritis over the past decade. Ann Rheum Dis 2015;74:738–45.

56. 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. Ann Rheum Dis 2010;69:995–1003.

57. Schipper LG, Vermeer M, Kuper HH, Hoekstra MO, Haagsma CJ, Den Broeder AA, et al. A tight control treatment strategy aiming for remission in early rheumatoid arthritis is more effective than usual care treatment in daily clinical practice: A study of two cohorts in the Dutch Rheumatoid Arthritis Monitoring registry. Ann Rheum Dis 2012;

58. Smolen JS, Landewé R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Ann Rheum Dis 2017;

59. Singh J, Furst D, Bharat A. 2012 Update of the 2008 American College of Rheumatology recommendations for the use of disease--modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:625–39.

60. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target : 2014 update of the recommendations of an international task force. 2015;1–13.

61. Kievit W, Fransen J, Oerlemans AJM, Kuper HH, van der Laar MAFJ, de Rooij DJRAM, et al. The efficacy of anti-TNF in rheumatoid arthritis, a comparison between randomised controlled trials and clinical practice. Ann Rheum Dis 2007;66:1473–8.

62. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. Value in health 2003;6:9–17.

63. Vermeer M, Kuper HH, Moens HJB, Drossaers-Bakker KW, Van Der Bijl AE, Van Riel PLCM, et al. Sustained beneficial effects of a protocolized treat-to-target strategy in very early rheumatoid arthritis: Three-year results of the dutch rheumatoid arthritis monitoring remission induction cohort. Arthritis Care Res (Hoboken) 2013;65:1219–26.

64. Steunebrink LMM, Vonkeman HE, ten Klooster PM, Hoekstra M, van Riel PLCM, van de Laar MAFJ. Recently diagnosed rheumatoid arthritis patients benefit from a treat-to-target strategy: results from the DREAM registry. Clin Rheumatol 2016;35:609–15.

65. Vermeer M, Kuper HH, Bernelot Moens HJ, Hoekstra M, Posthumus MD, van Riel PL, et al. Adherence to a treat-to-target strategy in early rheumatoid arthritis: results of the DREAM remission induction cohort. Arthritis Res Ther 2012;14:R254.

66. Kullback S, Kupperman M, Ku HH. Tests for Contingency Tables and Marltov Chains. Technometrics 1962;4:573–608.

67. Rubinstein RY, Kroese DP. Simulation and the Monte Carlo Method. Wiley 2008;377.

68. Jung J. Estimating Markov Transition Probabilities between Health States in the HRS Dataset. Indiana University 2006;1–42.

69. Rabin R, De Charro F. EQ-5D: A measure of health status from the EuroQol Group. In: Annals of Medicine. 2001.

70. Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, van Busschbach JJ. [Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff]. Ned Tijdschr Geneeskd 2005;

71. Hakkaart-van Roijen L, van der Linden N, Bouwmans C, Kanters T, Tan S. Kostenhandleiding. Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. 2015.

72. van der Ploeg, C.P.B., Pal, S., Vroome, E.D., Bossche, S. De kosten van ziekteverzuim voor werkgevers in nederland. 2014.

73. Schipper LG, Kievit W, den Broeder AA, van der Laar MA, Adang EMM, Fransen J, et al. Treatment strategies aiming at remission in early rheumatoid arthritis patients: Starting with methotrexate monotherapy is cost-effective. Rheumatology 2011;50:1320–30.

74. 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 Care Res (Hoboken) 2004;

75. Hallinen TA, Soini EJO, Eklund K, Puolakka K. Cost-utility of different treatment strategies after the failure of tumour necrosis factor inhibitor in rheumatoid arthritis in the Finnish setting. Rheumatology 2010;49:767–77.

76. Prevoo MLL, Vanthof MA, Kuper HH, Vanleeuwen MA, Vandeputte LBA, Vanriel P. Modified disease-activity scores that include 28-joint counts-development and validation in a prospective longitudinal-study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38:44–8.

77. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: A report of the ISPOR-SMDM modeling good research practices task force-6. Value in Health 2012;15:835–42.

78. Briggs A. Probabilistic analysis of cost-effectiveness models: Statistical representation of parameter uncertainty. Value in Health 2005;

79. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. A practical approach. Medical decision making 1985;5:157–77.

80. Steunebrink LMM, Versteeg GA, Vonkeman HE, Ten Klooster PM, Kuper HH, Zijlstra TR, et al. Initial combination therapy versus step-up therapy in treatment to the target of remission in daily clinical practice in early rheumatoid arthritis patients: results from the DREAM registry. Arthritis Res Ther 2015;18:60.

81. Markusse IM, Akdemir G, Dirven L, Goekoop-Ruiterman YPM, Van Groenendael JHLM, Han KH, et al. Long-term outcomes of patients with recent-onset rheumatoid arthritis after 10 years of tight controlled treatment: A randomized trial. Ann Intern Med 2016;164:523–31.

82. Combe B, Rincheval N, Benessiano J, Berenbaum F, Cantagrel A, Daurès JP, et al. Five-year favorable outcome of patients with early rheumatoid arthritis in the 2000s: Data from the ESPOIR cohort. Journal of Rheumatology 2013;40:1650–7.

83. Andersson MLE, Forslind K, Hafström I. Comparing Five Year Out-Come in Two Cohorts of Patients with Early Rheumatoid Arthritis – A BARFOT Study. Open Rheumatol J 2015;9:8–15.

84. Blom M, Kievit W, Donders ART, Den Broeder AA, Straten VHHP, Kuper I, et al. Effectiveness of a third tumor necrosis factor-α-blocking agent compared with rituximab after failure of 2 TNF-blocking agents in rheumatoid arthritis. Journal of Rheumatology 2011;38:2355–61.

85. Van Den Broek M, Lems WF, Allaart CF. BeSt practice: The success of early-targeted treatment in rheumatoid arthritis. Clin Exp Rheumatol 2012;30.

86. Mohd Jamid N, Shahril NS, Rajalingam S, Hussein H. AB0298 Treat to Target (T2T) in Rheumatoid Arthritis – a Single Center Analysis. Ann Rheum Dis 2015;74:993.1-993.

87. Rantalaiho V, Korpela M, Laasonen L, Kautiainen H, Järvenpää S, Hannonen P, et al. Early combination disease-modifying antirheumatic drug therapy and tight disease control improve long-term radiologic outcome in patients with early rheumatoid arthritis: the 11-year results of the Finnish Rheumatoid Arthritis Combination Therapy trial. Arthritis Res Ther 2010;12:R122.

88. Rémuzat C, Urbinati D, Mzoughi O, El Hammi E, Belgaied W, Toumi M. Overview of external reference pricing systems in Europe. J Mark Access Health Policy 2015;

89. Smolen JS. Treat-to-target: rationale and strategies. Clin Exp Rheumatol 30:S2-6.

90. Singh JA, Saag KG, Bridges SLJ, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. Arthritis Care Res (Hoboken) 2016;

91. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ 2006;15:1295–310.

92. Joensuu JT, Huoponen S, Aaltonen KJ, Konttinen YT, Nordström D, Blom M. The cost-effectiveness of biologics for the treatment of rheumatoid arthritis: A systematic review. PLoS One2015;

93. Taylor PC, Keystone EC, van der Heijde D, Weinblatt ME, del Carmen Morales L, Reyes Gonzaga J, et al. Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis. New England Journal of Medicine 2017;376:652–62.

94. Vermeer M, Kuper HH, Hoekstra M, Haagsma CJ, Posthumus MD, Brus HLM, et al. Implementation of a treat-to-target strategy in very early rheumatoid arthritis: results of the Dutch Rheumatoid Arthritis Monitoring remission induction cohort study. Arthritis Rheum 2011;63:2865–72.

95. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. 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;38:44–8.

96. Welsing PMJ, Severens JL, Hartman M, van Gestel AM, van Riel PLCM, Laan RFJM. The initial validation of a Markov model for the economic evaluation of (new) treatments for rheumatoid arthritis. Pharmacoeconomics 2006;24:1011–20.

97. Dougados M, Van Der Heijde D, Chen YC, Greenwald M, Drescher E, Liu J, et al. Baricitinib in patients with inadequate response or intolerance to conventional synthetic DMARDs: Results from the RA-BUILD study. Ann Rheum Dis 2017;

98. Tosh J, Stevenson M, Akehurst R. Health Economic Modelling of Treatment Sequences for Rheumatoid Arthritis: A Systematic Review. Curr Rheumatol Rep2014;

99. Zink A, Strangfeld A, Schneider M, Herzer P, Hierse F, Stoyanova-Scholz M, et al. Effectiveness of tumor necrosis factor inhibitors in rheumatoid arthritis in an observational cohort study: Comparison of patients according to their eligibility for major randomized clinical trials. Arthritis Rheum 2006;

100. van der Heijde DMFM, van Riel PLCM, van Rijswijk MH, van de Putte LBA. Influence of prognostic features on the final outcome in rheumatoid arthritis: A review of the literature. Semin Arthritis Rheum1988;

101. Smith E, Hoy D, Cross M, Merriman TR, Vos T, Buchbinder R, et al. The global burden of gout: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73:1470–6.

102. Becker MA, Schumacher HR, Benjamin KL, Gorevic P, Greenwald M, Fessel J, et al. Quality of life and disability in patients with treatment-failure gout. Journal of Rheumatology 2009;

103. Edwards NL, Sundy JS, Forsythe A, Blume S, Pan F, Becker MA. Work productivity loss due to flares in patients with chronic gout refractory to conventional therapy. J Med Econ 2011;

104. Bursill D, Taylor WJ, Terkeltaub R, Kuwabara M, Merriman TR, Grainger R, et al. Gout, Hyperuricemia, and Crystal-Associated Disease Network Consensus Statement Regarding Labels and Definitions for Disease Elements in Gout. Arthritis Care Res (Hoboken) 2019;71:427–34.

105. Khanna D, Fitzgerald JD, Khanna PP, Bae S, Singh MK, Neogi T, et al. 2012 American college of rheumatology guidelines for management of gout. part 1: Systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. Arthritis Care Res (Hoboken) 2012;

106. Richette P, Doherty M, Pascual E, Barskova V, Becce F, Castañeda-Sanabria J, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis2017;

107. Pascart T, Richette P. Current and future therapies for gout. Expert Opin Pharmacother2017;

108. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. Arthritis Rheum 2008;

109. Gandhi, PK, Gentry WM, Ma Q, Bottorff MB. Cost-Effectiveness Analysis of Allopurinol Versus Febuxostat in Chronic Gout Patients: A U.S. Payer Perspective. J Manag Care Spec Pharm 2016;

110. Smolen LJ, Gahn JC, Mitri G, Shiozawa A. Febuxostat in the management of gout: a cost-effectiveness analysis. J Med Econ 2015;

111. Jutkowitz E, Choi HK, Pizzi LT, Kuntz KM. Cost-effectiveness of allopurinol and febuxostat for the management of gout. Ann Intern Med 2014;161:617–26.

112. Janssens HJ, Janssen M, van de Lisdonk EH, van Riel PL, van Weel C. Use of oral prednisolone or naproxen for the treatment of gout arthritis: a double-blind, randomised equivalence trial. The Lancet 2008;

113. Saase L, Zwaap J, Knies S, Meijden van der C, Staal P, Heiden van der L. Cost-effectiveness in daily practice. 2015;64.

114. Kamatani N, Fujimori S, Hada T, Hosoya T, Kohri K, Nakamura T, et al. An Allopurinol-Controlled, Randomized, Double-Dummy, Double-Blind, Parallel Between-Group, Comparative Study of Febuxostat (TMX-67), a NonYPurine-Selective Inhibitor of Xanthine Oxidase, in Patients With Hyperuricemia Including Those With Gout in Japan. P. Journal of Clinical Rheumatology 2011;17:S13–8.

115. Huang X, Du H, Gu J, Zhao D, Jiang L, Li X, et al. An allopurinol-controlled, multicenter, randomized, double-blind, parallel between-group, comparative study of febuxostat in Chinese patients with gout and hyperuricemia. Int J Rheum Dis 2014;17:679–86.

116. Kamatani N, Fujimori S, Hada T, Hosoya T, Kohri K, Nakamura T, et al. An allopurinol-controlled, multicenter, randomized, open-label, parallel between-group, comparative study of febuxostat (TMX-67), a non-purine-selective inhibitor of xanthine oxidase, in patients with hyperuricemia including those with gout in Japan. Journal of Clinical Rheumatology 2011;17:S44–9.

117. Xu S, Liu X, Ming J, Chen S, Wang Y, Liu H, et al. A phase 3, multicenter, randomized, allopurinol-controlled study assessing the safety and efficacy of oral febuxostat in Chinese gout patients with hyperuricemia. 2015;18:669–78.

118. Kamatani N, Fujimori S, Hada T, Hosoya T, Kohri K, Nakamura T, et al. Placebo-Controlled Double-Blind Dose-Response Study of the Non-Purine-Selective Xanthine Oxidase Inhibitor Febuxostat (TMX-67) in Patients With Hyperuricemia (Including Gout Patients) in Japan. Journal of Clinical Rheumatology 2011;17:S35–43.

119. Becker MA, Schumacher HR, Espinoza LR, Wells AF, MacDonald P, Lloyd E, et al. The urate-lowering efficacy and safety of febuxostat in the treatment of the hyperuricemia of gout: the CONFIRMS trial. Arthritis Res Ther 2010;12:R63.

120. Becker MA, Schumacher HR, Wortmann RL, MacDonald PA, Palo WA, Eustace D, et al. Febuxostat, a novel nonpurine selective inhibitor of xanthine oxidase: A twenty-eight-day,

multicenter, phase II, randomized, double-blind, placebo-controlled, dose-response clinical trial examining safety and efficacy in patients with gout. Arthritis Rheum 2005;

121. Li S, Yang H, Guo Y, Wei F, Yang X, Li D, et al. Comparative efficacy and safety of urate-lowering therapy for the treatment of hyperuricemia: A systematic review and network meta-analysis. Sci Rep2016;

122. Zorginstituut Nederland. Drug prescription costs. [Internet. Accessed SEPTEMBER 27th, 2018. Website in Dutch] Available from: www.medicijnkosten.nl.

123. Sarawate CA, Patel PA, Schumacher HR, Yang W, Brewer KK, Bakst AW. Serum urate levels and gout flares: Analysis from managed care data. Journal of Clinical Rheumatology 2006;

124. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. BMJ 2003;326:472.

125. Briggs AH, Ades AE, Price MJ. Probabilistic sensitivity analysis for decision trees with multiple branches: Use of the Dirichlet distribution in a Bayesian framework. Medical Decision Making2003;

126. Smith E, Hoy D, Cross M, Merriman TR, Vos T, Buchbinder R, et al. The global burden of gout: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73:1470–6.

127. Bardin T, Voshaar MAHO, van de Laar MAFJ. The Human and Economic Burden of Difficult-to-Treat Gouty Arthritis. Joint Bone Spine 2015;

128. Lindemark F, Norheim OF, Johansson KA. Making use of equity sensitive QALYs: A case study on identifying the worse off across diseases. Cost Effectiveness and Resource Allocation 2014;

129. Chandratre P, Roddy E, Clarson L, Richardson J, Hider SL, Mallen CD. Health-related quality of life in gout: a systematic review. Rheumatology 2013;52:2031–40.

130. Stevenson M, Pandor A. Febuxostat for the management of hyperuricaemia in patients with gout: A nice single technology appraisal. Pharmacoeconomics 2011;

131. Perez-Ruiz F, Díaz-Torné C, Carcedo D. Cost-effectiveness analysis of febuxostat in patients with gout in Spain. J Med Econ 2016;19:604–10.

132. Beard SM, Von Scheele BG, Nuki G, Pearson I V. Cost-effectiveness of febuxostat in chronic gout. European Journal of Health Economics 2014;15:453–63.

133. Ruggeri M, Basile M, Drago C, Rolli FR, Cicchetti A. Cost-Effectiveness Analysis of Lesinurad/Allopurinol Versus Febuxostat for the Management of Gout/Hyperuricemia in Italy. Pharmacoeconomics 2018;

134. Sarawate CA, Brewer KK, Yang W, Patel PA, Schumacher HR, Saag KG, et al. Gout medication treatment patterns and adherence to standards of care from a managed care perspective. Mayo Clin Proc 2006;

135. Lesinurad/Allopurinol (Duzallo) for Gout-Associated HyperuricemiaLesinurad/Allopurinol (Duzallo) for Gout-Associated HyperuricemiaLesinurad/Allopurinol (Duzallo) for Gout-Associated Hyperuricemia. JAMA 2018;319:188–9.

136. Wells GA, Russell AS, Haraoui B, Bissonnette R, Ware CF. Validity of quality of life measurement tools—from generic to disease-specific. J Rheumatol Suppl 2011;88:2–6.

137. Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic and specific quality-of-life instruments. J Clin Epidemiol 2003;56:52–60.

138. Schett G, Emery P, Tanaka Y, Burmester G, Pisetsky DS, Naredo E, et al. Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. Ann Rheum Dis 2016;75:1428–37.

139. Eland-de Kok P, van Os-Medendorp H, Vergouwe-Meijer A, Bruijnzeel-Koomen C, Ros W. A systematic review of the effects of e-health on chronically ill patients. J Clin Nurs 2011;20:2997–3010.

140. Katz JN, Phillips CB, Fossel AH, Liang MH. Stability and responsiveness of utility measures. Med Care 1994;32:183–8.

141. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40:171–8.

142. Harrison MJ, Davies LM, Bansback NJ, McCoy MJ, Verstappen SMM, Watson K, et al. The comparative responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory arthritis. Quality of Life Research 2009;18:1195–205.

143. Linde L, Sørensen J, Østergaard M, Hørslev-Petersen K, Hetland ML. Health-related quality of life: validity, reliability, and responsiveness of SF-36, EQ-15D, EQ-5D, RAQoL, and HAQ in patients with rheumatoid arthritis. J Rheumatol 2008;35:1528–37.

144. Sonntag M, Konnopka A, Leichsenring F, Salzer S, Beutel ME, Herpertz S, et al. Reliability, validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with social phobia. Health Qual Life Outcomes 2013;11:1–9.

145. Payakachat N, Ali MM, Tilford JM. Can the EQ-5D detect meaningful change? A systematic review. Pharmacoeconomics 2015;33:1137–54.

146. Peasgood T, Brazier J, Papaioannou D. A systematic review of the validity and responsiveness of EQ-5D and SF-6D for depression and anxiety. HEDS Discussion paper 12/15 2012;

147. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102–9.

148. Ghiti Moghadam M, Lamers-Karnebeek FBG, Vonkeman HE, Ten Klooster PM, Tekstra J, Schilder AM, et al. Multi-biomarker disease activity score as a predictor of disease relapse

in patients with rheumatoid arthritis stopping TNF inhibitor treatment. PLoS One 2018;13:e0192425.

149. Isik A, Koca SS, Ozturk A, Mermi O. Anxiety and depression in patients with rheumatoid arthritis. Clin Rheumatol 2007;26:872–8.

150. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, et al. Stopping Tumor Necrosis Factor Inhibitor Treatment in Patients With Established Rheumatoid Arthritis in Remission or With Stable Low Disease Activity: A Pragmatic Multicenter, Open-Label Randomized Controlled Trial. Arthritis Rheumatol 2016;68:1810–7.

151. Ghiti Moghadam M, ten Klooster PM, Vonkeman HE, Kneepkens EL, Klaasen R, Stolk JN, et al. Impact of Stopping Tumor Necrosis Factor-inhibitors on Rheumatoid Arthritis Patients' Burden of Disease. Arthritis Care Res (Hoboken) 2018;70:516–24.

152. van der Maas A, Lie E, Christensen R, Choy E, de Man YA, van Riel P, et al. Construct and criterion validity of several proposed DAS28-based rheumatoid arthritis flare criteria: an OMERACT cohort validation study. Ann Rheum Dis 2013;72:1800–5.

153. Ghiti Moghadam M, Vonkeman HE, Ten Klooster PM, Tekstra J, van Schaardenburg D, Starmans-Kool M, et al. Stopping Tumor Necrosis Factor Inhibitor Treatment in Patients With Established Rheumatoid Arthritis in Remission or With Stable Low Disease Activity: A Pragmatic Multicenter, Open-Label Randomized Controlled Trial. Arthritis Rheumatol 2016;68:1810–7.

154. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. 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;38:44–8.

155. Ware Jr JE. Health Survey Update. Spine (Phila Pa 1976) 2000;25:3130–9.

156. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137–45.

157. The EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.

158. ten Klooster PM, Vonkeman HE, Taal E, Siemons L, Hendriks L, de Jong AJL, et al. Performance of the Dutch SF-36 version 2 as a measure of health-related quality of life in patients with rheumatoid arthritis. Health Qual Life Outcomes 2013;11:77.

159. Ware Jr JE. Health Survey Update. Spine (Phila Pa 1976) 2000;25:3130–9.

160. ten Klooster PM, Taal E, van de Laar MA. Rasch analysis of the Dutch Health Assessment Questionnaire disability index and the Health Assessment Questionnaire II in patients with rheumatoid arthritis. Arthritis Rheum 2008;59:1721–8.

161. The EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.

162. Lamers LM, McDonnell J, Stalmeier PFM, Krabbe PFM, Busschbach JJ V. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ 2006;15:1121–32.

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

164. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;632–42.

165. Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995;22:1191–2.

166. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459-68.

167. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis 1986;39:897-906.

168. Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd ed. New York, NY: John Wiley & Sons; 2000.

169. Hays RD, Revicki D. Reliability and validity (including responsiveness). In: Fayers PM, Hays RD, editors. Assessing quality of life in clinical trials: Methods and practice. Oxford: Oxford University Press; 2005. page 25–39.

170. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985;28:542–7.

171. van der Maas A, Lie E, Christensen R, Choy E, de Man YA, van Riel P, et al. Construct and criterion validity of several proposed DAS28-based rheumatoid arthritis flare criteria: an OMERACT cohort validation study. Ann Rheum Dis 2013;72:1800–5.

172. Ghiti Moghadam M, Lamers-Karnebeek FBG, Vonkeman HE, ten Klooster PM, Tekstra J, Schilder AM, et al. Multi-biomarker disease activity score as a predictor of disease relapse in patients with rheumatoid arthritis stopping TNF inhibitor treatment. PLoS One 2018;13:e0192425.

173. Linde L, Sørensen J, Østergaard M, Hørslev-Petersen K, Hetland ML. Health-related quality of life: validity, reliability, and responsiveness of SF-36, EQ-15D, EQ-5D, RAQoL, and HAQ in patients with rheumatoid arthritis. J Rheumatol 2008;35:1528–37.

174. Stark RG, Reitmeir P, Leidl R, König HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. Inflamm Bowel Dis 2010;16:42–51.

175. Obradovic M, Lal A, Liedgens H. Validity and responsiveness of EuroQol-5 dimension (EQ-5D) versus Short Form-6 dimension (SF-6D) questionnaire in chronic pain. Health Qual Life Outcomes 2013;11:1–9.

176. Buitinga L, Braakman-Jansen LMA, Taal E, Kievit W, Visser H, van Riel PLCM, et al. Comparative responsiveness of the EuroQol-5D and Short Form 6D to improvement in patients with rheumatoid arthritis treated with tumor necrosis factor blockers: Results of the Dutch Rheumatoid Arthritis Monitoring registry. Arthritis Care Res (Hoboken) 2012;64:826–32.

177. Veehof MM, ten Klooster PM, Taal E, van Riel PLCM, van de Laar MAFJ. Comparison of internal and external responsiveness of the generic medical outcome study short form-36 (SF-36) with disease-specific measures in rheumatoid arthritis. Journal of Rheumatology 2008;35:610–7.

178. Katz N. The impact of pain management on quality of life. J Pain Symptom Manage 2002;24:S38–47.

179. Burmester GR, Pope JE. Novel treatment strategies in rheumatoid arthritis. The Lancet 2017;389:2338–48.

180. Smolen JS, Landewe R, Bijlsma J, Burmester G, Chatzidionysiou K, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. Ann Rheum Dis 2017;

181. Taylor PC, Keystone EC, van der Heijde D, Weinblatt ME, del Carmen Morales L, Reyes Gonzaga J, et al. Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis. New England Journal of Medicine 2017;376:652–62.

182. Van Vollenhoven RF, Fleischmann R, Cohen S, Lee EB, García Meijide JA, Wagner S, et al. Tofacitinib or adalimumab versus placebo in rheumatoid arthritis. New England Journal of Medicine 2012;367:508–19.

183. Combe B, Kivitz A, Tanaka Y, van Der Heijde D, Simon JA, Baraf HSB, et al. Filgotinib versus placebo or adalimumab in patients with rheumatoid arthritis and inadequate response to methotrexate: a phase III randomised clinical trial. Ann Rheum Dis 2021;80:848–58.

184. Fleischmann R, Pangan AL, Song I, Mysler E, Bessette L, Peterfy C, et al. Upadacitinib versus placebo or adalimumab in patients with rheumatoid arthritis and an inadequate response to methotrexate: results of a phase III, double-blind, randomized controlled trial. Arthritis & Rheumatology 2019;71:1788–800.

185. Fleischmann R, Schiff M, van der Heijde D, Ramos-Remus C, Spindler A, Stanislav M, et al. Baricitinib, methotrexate, or combination in patients with rheumatoid arthritis and no or limited prior disease-modifying antirheumatic drug treatment. Arthritis & Rheumatology 2017;69:506–17.

186. Avci AB, Feist E, Burmester GR. Early phase studies of JAK1 selective inhibitors in rheumatoid arthritis. Rheumatology (United Kingdom) 2021;60:II11–6.

187. Alten R, Krüger K, Rellecke J, Schiffner-Rohe J, Behmer O, Schiffhorst G, et al. Examining patient preferences in the treatment of rheumatoid arthritis using a discrete-choice approach. Patient Prefer Adherence 2016;2217–28.

188. Smolen JS, Landewé RBM, Bijlsma JWJ, Burmester GR, Dougados M, Kerschbaumer A, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020;79:685–99.

189. Nash P, Kerschbaumer A, Dörner T, Dougados M, Fleischmann RM, Geissler K, et al. Points to consider for the treatment of immune-mediated inflammatory diseases with Janus kinase inhibitors: a consensus statement. Ann Rheum Dis [Internet] 2021;80:71 LP – 87. Available from: http://ard.bmj.com/content/80/1/71.abstract

190. Ford I, Norrie J. Pragmatic trials. New England journal of medicine 2016;375:454–63.

191. Smolen JS, Landewé RBM, Bergstra SA, Kerschbaumer A, Sepriano A, Aletaha D, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2022 update. Ann Rheum Dis 2023;82:3–18.

192. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology 1995;38:727–35.

193. Pincus T, Swearingen CJ, Bergman M, Yazici Y. RAPID3 (Routine Assessment of Patient Index Data 3), a rheumatoid arthritis index without formal joint counts for routine care: proposed severity categories compared to disease activity score and clinical disease activity index categories. J Rheumatol 2008;35:2136–47.

194. Van Riel PL. The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28). Clin Exp Rheumatol 2014;32:S65-74.

195. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology 2003;42:244–57.

196. Van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol 2000;27:261–3.

197. MDCG. MDCG 2021-6 Regulation (EU) 2017 / 745 – Questions & Answers regarding clinical investigation. Medical Device Coordination Group 2021;1–19.

198. Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. Trials 2011;12:1–12.

199. Aaltonen KJ, Ylikylä S, Tuulikki Joensuu J, Isomäki P, Pirilä L, Kauppi M, et al. Efficacy and effectiveness of tumour necrosis factor inhibitors in the treatment of rheumatoid arthritis in randomized controlled trials and routine clinical practice. Rheumatology 2017;56:725–35.

200. Zink A, Strangfeld A, Schneider M, Herzer P, Hierse F, Stoyanova-Scholz M, et al. Effectiveness of tumor necrosis factor inhibitors in rheumatoid arthritis in an observational cohort study: Comparison of patients according to their eligibility for major randomized clinical trials. Arthritis Rheum 2006;

201. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj 2009;338.

202. Katchamart W, Johnson S, Lin HL, Phumethum V, Salliot C, Bombardier C. Predictors for remission in rheumatoid arthritis patients: a systematic review. Arthritis Care Res (Hoboken) 2010;62:1128–43.

203. Hamann PDH, Pauling JD, McHugh N, Shaddick G, Hyrich K. Predictors, demographics and frequency of sustained remission and low disease activity in anti-tumour necrosis factor–treated rheumatoid arthritis patients. Rheumatology 2019;58:2162–9.

204. Hamann P, Holland R, Hyrich K, Pauling JD, Shaddick G, Nightingale A, et al. Factors associated with sustained remission in rheumatoid arthritis in patients treated with anti–tumor necrosis factor. Arthritis Care Res (Hoboken) 2017;69:783–93.

205. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 1998;17:857–72.

206. Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. Stat Methods Med Res 2015;24:224–54.

207. Burmester GR, Rubbert-Roth A, Cantagrel A, Hall S, Leszczynski P, Feldman D, et al. A randomised, double-blind, parallel-group study of the safety and efficacy of subcutaneous tocilizumab versus intravenous tocilizumab in combination with traditional disease-modifying antirheumatic drugs in patients with moderate to severe rheumatoid art. Ann Rheum Dis 2014;73:69–74.

208. Weinblatt ME, Schiff M, Valente R, van der Heijde D, Citera G, Zhao C, et al. Head-to-head comparison of subcutaneous abatacept versus adalimumab for rheumatoid arthritis: findings of a phase IIIb, multinational, prospective, randomized study. Arthritis Rheum 2013;65:28–38.

209. McKnight PE, Najab J. Mann-Whitney U Test. The Corsini encyclopedia of psychology 2010;1.

210. Landewé R, van der Heijde D. Radiographic progression depicted by probability plots: presenting data with optimal use of individual values. Arthritis Rheum 2004;50:699–706.

211. Cuschieri S. The CONSORT statement. Saudi J Anaesth 2019;13:S27.

212. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. Bmj 2008;337.

213. Venetsanopoulou AI, Voulgari P V, Drosos AA. Janus kinase versus TNF inhibitors: Where we stand today in rheumatoid arthritis. Expert Rev Clin Immunol 2022;18:485–93.

214. Rantalaiho V, Kautiainen H, Korpela M, Hannonen P, Kaipiainen-Seppänen O, Möttönen 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 NE. Ann Rheum Dis 2014;73:1954–61.

215. Kuusalo L, Puolakka K, Kautiainen H, Blåfield H, Eklund KK, Ilva K, et al. Impact of physicians adherence to treat-to-target strategy on outcomes in early rheumatoid arthritis in the NEO-RACo trial. Scand J Rheumatol 2015;44:449–55.

216. Grigor C, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, et al. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. The Lancet 2004;364:263–9.

217. Schipper LG, Vermeer M, Kuper HH, Hoekstra MO, Haagsma CJ, Den Broeder AA, et al. A tight control treatment strategy aiming for remission in early rheumatoid arthritis is more effective than usual care treatment in daily clinical practice: A study of two cohorts in the Dutch Rheumatoid Arthritis Monitoring registry. Ann Rheum Dis 2012;

218. Steunebrink LMM, Vonkeman HE, ten Klooster PM, Hoekstra M, van Riel PLCM, van de Laar MAFJ. Recently diagnosed rheumatoid arthritis patients benefit from a treat-to-target strategy: results from the DREAM registry. Clin Rheumatol 2016;35:609–15.

219. Fleischmann R, Mysler E, Hall S, Kivitz AJ, Moots RJ, Luo Z, et al. Efficacy and safety of tofacitinib monotherapy, tofacitinib with methotrexate, and adalimumab with methotrexate in patients with rheumatoid arthritis (ORAL Strategy): a phase 3b/4, double-blind, head-to-head, randomised controlled trial. The Lancet 2017;390:457–68.

220. Westhovens R, Rigby WFC, Van Der Heijde D, Ching DWT, Stohl W, Kay J, et al. Filgotinib in combination with methotrexate or as monotherapy versus methotrexate monotherapy in patients with active rheumatoid arthritis and limited or no prior exposure to methotrexate: The phase 3, randomised controlled FINCH 3 trial. Ann Rheum Dis 2021;80:727–38.

221. Rubbert-Roth A, Enejosa J, Pangan AL, Haraoui B, Rischmueller M, Khan N, et al. Trial of upadacitinib or abatacept in rheumatoid arthritis. New England Journal of Medicine 2020;383:1511–21.

222. Fleischmann RM, Blanco R, Hall S, Thomson GTD, Van Den Bosch FE, Zerbini C, et al. Switching between janus kinase inhibitor upadacitinib and adalimumab following

insufficient response: Efficacy and safety in patients with rheumatoid arthritis. Ann Rheum Dis 2021;80:432–9.

223. Siemons L, Vonkeman HE, Ten Klooster PM, Van Riel PLCM, Van De Laar MAFJ. Interchangeability of 28-joint disease activity scores using the erythrocyte sedimentation rate or the C-reactive protein as inflammatory marker. Clin Rheumatol 2014;33:783–9.

224. Fleischmann RM, van der Heijde D, Gardiner P V, Szumski A, Marshall L, Bananis E. DAS28-CRP and DAS28-ESR cut-offs for high disease activity in rheumatoid arthritis are not interchangeable. RMD Open 2017;3:e000382.

225. Harrington R, Al Nokhatha SA, Conway R. JAK inhibitors in rheumatoid arthritis: an evidence-based review on the emerging clinical data. J Inflamm Res 2020;519–31.

226. EMA. EMA recommends measures to minimise risk of serious side effects with Janus kinase inhibitors for chronic inflammatory disorders [Internet]. EUROPEAN MEDICINES AGENCY website [cited 2023 Aug 4];Available from: https://www.ema.europa.eu/en/news/ema-recommends-measures-minimise-risk-serious-side-effects-janus-kinase-inhibitors-chronic

227. Rosenblum H, Amital H. Anti-TNF therapy: safety aspects of taking the risk. Autoimmun Rev 2011;10:563–8.

228. Dixon WG, Hyrich KL, Watson KD, Lunt M, Galloway J, Ustianowski A, et al. Drug-specific risk of tuberculosis in patients with rheumatoid arthritis treated with anti-TNF therapy: results from the British Society for Rheumatology Biologics Register (BSRBR). Ann Rheum Dis 2010;69:522–8.

229. Hussain A, Tarahomi T, Singh L, Bollampally M, Heydari-Kamjani M, Kesselman MM. Cardiovascular Risk Associated With TNF Alpha Inhibitor Use in Patients With Rheumatoid Arthritis. Cureus 2021;13.

230. Ministerie van Volksgezondheid W en S, Dijkstra P. Voortgangsbrief financiële arrangementen geneesmiddelen 2023 [Internet]. Den Haag: 2024 [cited 2024 Dec 7]. Available from: https://www.rijksoverheid.nl/documenten/kamerstukken/2024/06/13/kamerbrief-over-voortgangsbrief-financiele-arrangementen-geneesmiddelen-2023

231. Commissie Richtlijnherziening. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidzorg [Internet]. 2024 [cited 2024 Dec 7]. Available from: https://www.zorginstituutnederland.nl/over-ons/publicaties/publicatie/2024/01/16/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg

PhD Portfolio

International conferences

American College of Rheumatology ,San Diego 2017

- poster presentation

American College of Rheumatolgoy, Atlanta 2019

- oral presentation

American College of Rheumatology, San Diego 2023

European League Against Rheumatism, Madrid 2017

- oral presentation

European League Against Rheumatism, Amsterdam 2022

European League against Rheumatism, Milan 2023

National conferences

Nederlandse Vereniging Reumatologie, NVR, Arnhem 2017

- poster presentation

Courses and Trainings

Menzis Talent Ontwikkel Programma	2022-2023
Inzicht In Invloed Bureau Zuidema	2022
NFU eBROK	2020
TreeAge Software Healthcare Advanced Markov/DES training	2017
Journal Club Transparency in Healthcare	2017-2023

About the author

Celine van de Laar is a health economist. She obtained a BSc in Economics from the University of Groningen in 2017. She obtained a MSc of Economics, specialising in Health Economics at Erasmus University Rotterdam in 2019. She has worked on her research ever since as an external PhD candidate at the Erasmus School of Health Policy and Management. Her research interests lie at the intersection of clinical medicine research and health economics.