PharmacoEconomics (2014) 32:1015–1028
DOI 10.1007/s40273-014-0184-4

ORIGINAL RESEARCH ARTICLE

Modelling Outcomes of Complex Treatment Strategies Following a Clinical Guideline for Treatment Decisions in Patients with Rheumatoid Arthritis

An Tran-Duy · Annelies Boonen · Wietske Kievit · Piet L. C. M. van Riel · Mart A. F. J. van de Laar · Johan L. Severens

Published online: 28 June 2014

Abstract

Background Management of rheumatoid arthritis (RA) is characterised by a sequence of disease-modifying anti-rheumatic drugs (DMARDs) and biological response modifiers (BRMs). In most of the Western countries, the drug sequences are determined based on disease activity and treatment history of the patients. A model for realistic patient outcomes should reflect the treatment pathways relevant for patients with specific characteristics.

Objective This study aimed at developing a model that could simulate long-term patient outcomes and cost effectiveness of treatment strategies with and without inclusion of BRMs following a clinical guideline for treatment decisions.

Methods Discrete event simulation taking into account patient characteristics and treatment history was used for model development. Treatment effect on disease activity, costs, health utilities and times to events were estimated using Dutch observational studies. Long-term progression of physical functioning was quantified using a linear mixed-effects model. Costs and health utilities were estimated using two-part models. The treatment strategy recommended by the Dutch Society for Rheumatology where both DMARDs and BRMs were available (Strategy 2) was compared with the treatment strategy without BRMs (Strategy 1). Ten thousand theoretical patients were tracked individually until death. In the probabilistic sensitivity analysis, Monte Carlo simulations were performed with 1,000 sets of parameters sampled from appropriate probability distributions.

Results The simulated changes over time in disease activity and physical functioning were plausible. The incremental cost per quality-adjusted life-year gained of Strategy 2 compared with Strategy 1 was €124,011. At a willingness-to-pay threshold higher than €119,167, Strategy 2 dominated Strategy 1 in terms of cost effectiveness but the probability that the Strategy 2 is cost effective never exceeded 0.87.

Conclusions It is possible to model the outcomes of complex treatment strategies based on a clinical guideline for the management of RA. Following the Dutch guideline and using real-life data, inclusion of BRMs in the treatment

Electronic supplementary material The online version of this article (doi:10.1007/s40273-014-0184-4) contains supplementary material, which is available to authorized users.

A. Tran-Duy (✉)
Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Center, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands
e-mail: an.tranduy@mumc.nl

A. Tran-Duy · A. Boonen
Division of Rheumatology, Department of Internal Medicine, Maastricht University Medical Center, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands

A. Tran-Duy · A. Boonen
Caphri School for Public Health and Primary Care, Maastricht UMC+, P.O. Box 616, 6200 MD Maastricht, The Netherlands

W. Kievit · P. L. C. M. van Riel
Department of Rheumatic Diseases, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

M. A. F. J. van de Laar
Department of Rheumatology and Clinical Immunology, University Twente & Medisch Spectrum Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

J. L. Severens
Institute of Health Policy and Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

strategy for RA appeared to be less favourable in our model than in most of the existing models that compared drug sequences independent of patient characteristics and used data from randomised controlled clinical trials. Despite complexity and demand for extensive data, our modelling approach can help to identify the knowledge gaps in clinical guidelines for RA management and priorities for future research.

Key Points for Decision Makers

Treatment strategies for patients with early-stage rheumatoid arthritis (RA) are becoming increasingly complex and it is therefore necessary to develop a modelling approach that fully supports economic evaluation of a wide range of real-life treatment strategies

Treatment strategies including biological response modifiers (BRMs) with treatment decisions based on patient characteristics and treatment history have been accepted in most of the Western countries, but few cost-effectiveness models have attempted to take the use of BRMs in clinical practice into consideration

Modelling the outcomes of complex treatment strategies based on a clinical guideline for the management of RA is possible and helpful in assessing the impact of changes in clinical practice on disease progression, health and costs of the RA patients, and cost effectiveness of the new treatment decisions

Following the Dutch guideline for management of RA and using real-life data, the simulated incremental cost per quality-adjusted life-year gained of the treatment strategy including BRMs compared with the treatment strategy without BRMs was higher than conventionally accepted willingness-to-pay thresholds, and the probability that the treatment strategy including BRMs is cost effective never exceeded 0.87

1 Introduction

In the management of rheumatoid arthritis (RA), early treatment with synthetic disease-modifying antirheumatic drugs (DMARDs) and subsequent use of biological response modifiers (BRMs) to achieve low disease activity or remission have been recommended for clinical practice in both American and European treatment guidelines [1–4]. Because BRMs are far more expensive than DMARDs, the incremental costs per quality-adjusted life-year (QALY) [5] gained (ICERs) of treatments with BRMs have been discussed extensively.

Despite the growing number of published studies on the cost effectiveness of BRMs, few have attempted to take the use of BRMs in clinical practice into consideration. The sequence of drugs in the real-life management of RA is determined based on characteristics, notably disease activity using the Disease Activity Score for 28 joints (DAS28) [3, 6, 7], and treatment history of the patients. It is well recognised that the effectiveness of different drugs and of the same drug at different positions in a drug sequence are most likely different. Additionally, early or late administration of a drug affects the course of disease and thus the health and costs of the patients. Therefore, a model for realistic patient outcomes should reflect the treatment pathways relevant for patients with specific characteristics. Recently, there has been an increasing tendency of adopting sequential treatment strategies instead of single drugs as comparators in the cost-effectiveness models for RA in different European countries [8–15]. However, the order of the drugs in the existing models was determined independently of disease activity and treatment history of the patients.

The present study aimed at developing a model that could simulate long-term patient outcomes and cost effectiveness of treatment strategies with and without inclusion of BRMs; following a formal clinical guideline and rheumatologist opinion for treatment decisions based on patient characteristics and treatment history, and using real-life data to parameterize the model.

2 Methods

2.1 Model Structure

Discrete event simulation (DES) was used to track changes over time in disease measures, utilities and costs of the individual patients. This modelling approach can overcome intractable problems in conventional models for a chronic disease (e.g. numerous chance nodes in a decision tree or a large number of states in a Markov model) [16–18]. Moreover, memorising changes in patient characteristics over time is possible in DES, which permits simulation of the rheumatologist decisions on starting and switching treatments based on patient history [19].

The main components of the DES model consist of entities, states and events. An *entity* is an RA patient with a set of characteristics (e.g. age, gender, rheumatic factor, disease duration, disease activity and physical functioning), which influence the simulation outcomes.

States in our model were defined based on trends of change in DAS28 and by on-going treatments to help determine potential events that may occur during tracking disease progression and treatment decisions. Different from those in a Markov model, these states are not necessarily mutually exclusive. Conceptually, a patient can be in three different disease activity phases (DAS28-related states) while treated with a specific drug [20]. First, the decreasing phase is characterized by a steady decrease in DAS28 as a response to the treatment. Second, the maintenance phase is characterized by small fluctuations of DAS28 surrounding a constant level. Third, the increasing phase is characterized by a continuous increase in DAS28 after the patient stops responding to the drug. In this case, DAS28 was assumed to return to the patient's baseline level within 12 weeks. When the patient is receiving 'palliative' therapy, i.e. the use of only low-dose corticosteroids after failure of all drugs in a treatment strategy, DAS28 was assumed to be stable. These DAS28-related states were used to determine the events that may occur given a trend of change in DAS28 (see Table 1).

Regarding the treatments, a patient can be in one of eight states, referred to as treatment-related states, which is being treated with the first or second DMARD, the first or second tumour necrosis factor α (TNF) inhibitor BRM, the first or second non-TNF BRM, the DMARDs after failure of the first two DMARDs, or 'palliative' therapy. Treatment-related states were used to determine changes in DAS28, times to DAS28-related events and a new treatment when the current drug fails.

Events in the model were classified as DAS28-related events and DAS28-neutral events. DAS28-related events consisted of "End of DAS28 decrease", "Loss of response to the current drug" (which caused an increase in DAS28) and "DAS28 reaching 1.2 unit higher". The last event was formulated to help calculate the rate of DAS28 increase based on the assumption that DAS28 returned to the baseline level in 12 weeks after a loss of response to the current treatment, and that a 1.2-unit change in DAS28 was significant [21]. DAS28-neutral events consisted of "Severe toxicity of the current drug", "Visit a rheumatologist", "Select a new treatment", "Start a new treatment" (i.e. first administration of a new treatment) and "Death". A summary of the interdependence among states and events is given in Table 1. For competing events, the patient will "jump" to the event to which the sampled time is shortest. When an event occurred, an associated procedure was invoked for implementation where the patient characteristics were updated and times to the next events were computed. It should be noted that "Select a new treatment" and "Start a new treatment" may immediately follow "Visit a rheumatologist", but "Visit to a rheumatologist" does not necessarily lead to the occurrence of "Select a new treatment" and therefore not "Start of a new treatment" as the next events. This is the reason for a separation of these three events. The general simulation process is depicted in Fig. 1. Details on the simulation process can be found in the online supplementary appendix.

2.2 Treatment Strategies

We compared the treatment strategy recommended by the Dutch Society for Rheumatology where both DMARDs and BRMs were available (Strategy 2) with the treatment strategy without inclusion of BRMs (Strategy 1). Availability and order of the drugs in these two strategies were as follows:

- Strategy 1: eight DMARDs available in the following sequence: methotrexate followed randomly by sulphasalazine or leflunomide, which was followed by azathioprine, cyclosporin A, cyclophosphamide, hydroxychloroquine and injectable gold at a random order;

Table 1 Interdependency of the model states and events

DAS28-related state	Possible DAS28-related event	Possible DAS28-neutral event	Possible treatment-related state
On DAS28 decrease	End of DAS28 decrease[a]	Visit a rheumatologist[a]	On DMARDs
On DAS28 maintenance	Loss of response[a]	Severe toxicity[b]	On anti-TNF BRMs
On DAS28 increase	DAS28 reaching 1.2 unit higher[a]	Select a new treatment[b]	On non-TNF BRMs
		Start a new treatment[b]	
		Death[a]	
On DAS28 stability	–	Visit a rheumatologist[a]	On 'palliative' therapy
		Death[a]	

DAS28 Disease Activity Score for 28 joints, *DMARD* disease-modifying antirheumatic drug, *TNF* tumour necrosis factor, *BRMs* biological response modifiers, *'palliative' therapy* the use of only low-dose corticosteroids after failure of all drugs in a treatment strategy

[a] Competing events

[b] Events that occur only when a visit to a rheumatologist occurs

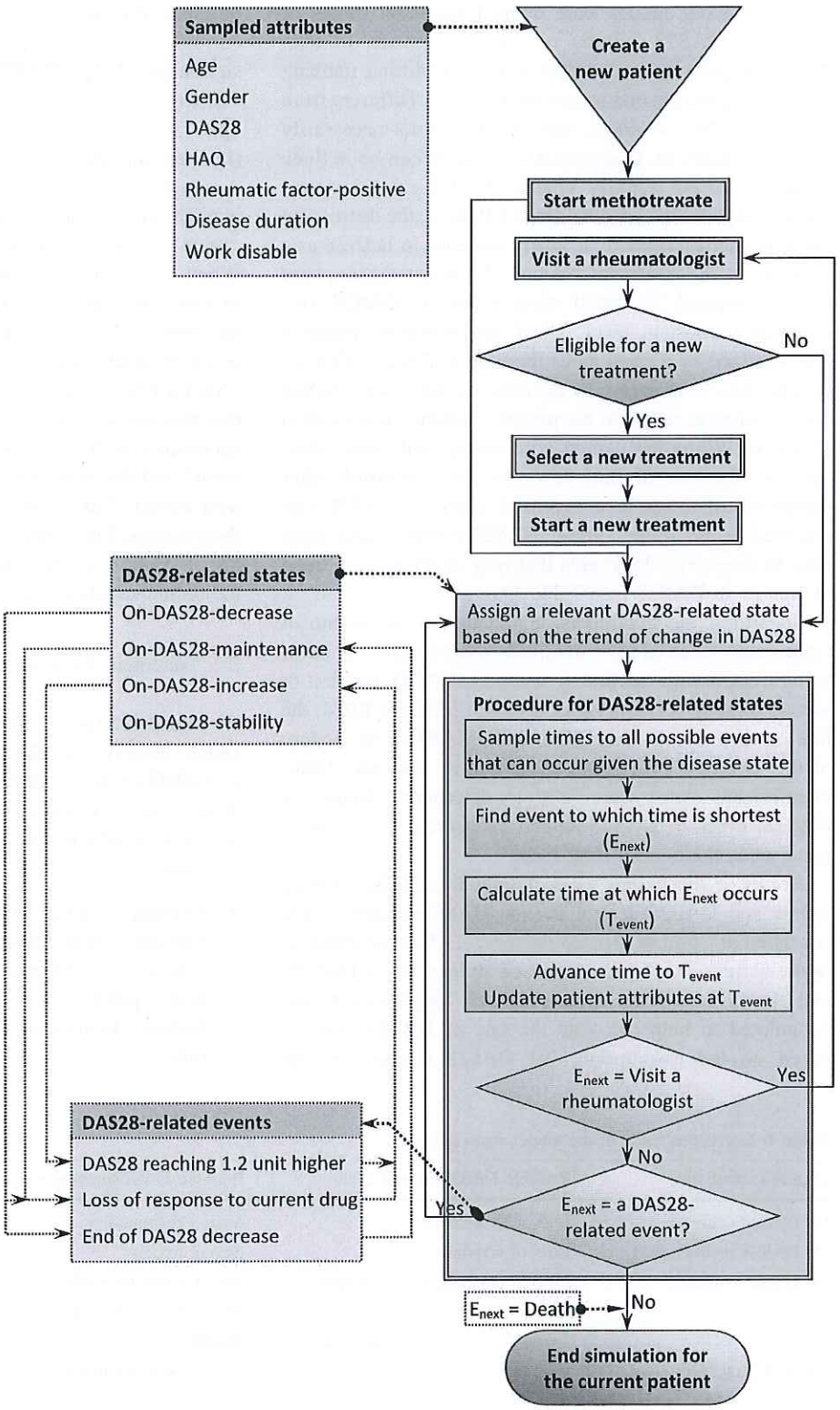

Fig. 1 Flow chart of the general simulation process. The *triangle at the top* represents the entry of the simulation. The *rectangles with double-line* and *single-line borders* represent procedures and instructions, respectively. The *diamonds* represent decisions. The *boxes with dotted borders* provide extra information for the flow chart. See text and the online supplementary appendix for details. *DAS28* Disease Activity Score for 28 joints, *HAQ* Health Questionnaire Assessment Disability Index

- Strategy 2: eight DMARDs as those in Strategy 1, and two anti-TNF and two non-TNF BRMs, available in the following sequence: methotrexate followed randomly by sulphasalazine or leflunomide, two anti-TNF BRMs, and two non-TNF BRMs, which were followed by azathioprine, cyclosporin A, cyclophosphamide,

hydroxychloroquine and injectable gold at a random order. The two anti-TNF BRMs comprised one monoclonal antibody and one soluble receptor fusion protein, which were randomly chosen from etanercept, adalimumab, infliximab, golimumab and certolizumab. The two non-TNF BRMs were randomly chosen from rituximab, abatacept and tocilizumab. For patients with a positive rheumatic factor, rituximab must be available at the first position of the non-TNF BRMs; for patients with a negative rheumatic factor, abatacept or tocilizumab must be available at the first position of the non-TNF BRMs.

The order of the drugs and random selections between sulphasalazine and leflunomide, among anti-TNF BRMs, among non-TNF BRMs and among DMARDs other than sulphasalazine and leflunomide in the above-mentioned treatment strategies agreed well with those recorded in the observational cohorts used for model inputs (see Sect. 2.3). The algorithm for treatment decisions in the model was formulated based on the Dutch guidelines and rheumatologist opinion for the treatment of RA. First, a flow chart of treatment decisions was drawn based on the general guidelines and delivered to the rheumatologists at the Universities of Maastricht, Nijmegen and Twente in The Netherlands. Second, several meetings were organized among the stakeholders to discuss and refine the flow chart based on personal experience and knowledge about daily clinical practice. Finally, the algorithm was concluded based on a consensus on the treatment decisions that reflect the predominant clinical practice of the management of RA in The Netherlands. In this algorithm, methotrexate, in combination with intramuscularly injected corticosteroids to bridge the time of action on disease activity of the first DMARD, was started as soon as RA is diagnosed. The next treatment based on the order of the drugs in the respective strategy was considered when a drug failed primarily or secondarily, or caused severe toxicity. Primary failure was assumed if DAS28 was still higher than 3.2 after 3–6 months (random) since the start of the treatment. Secondary failure was assumed when DAS28 went back to a level >3.2 after a primary response to the treatment. In Strategy 2, BRMs were combined with methotrexate if the patient did not experience severe toxicity when receiving methotrexate monotherapy after RA diagnosis. When there were no DMARDs left in the sequence, 'palliative' therapy was given in both strategies.

2.3 Model Inputs and Data Sources

2.3.1 Patients' Baseline Characteristics

The initial cohort to be simulated consisted of patients who were newly diagnosed with RA. We used the registry of the Radboud University Nijmegen Medical Centre (Nijmegen Inception Cohort) [22] to estimate the empirical distributions of baseline age and gender, DAS28 and the presence of positive rheumatic factor. This observational study was started in 1985 to collect information on the long-term course of RA in patients with disease duration <1 year and being DMARD naïve on the inclusion date. Because the course of disease in RA patients has become milder in recent years [23], only data of patients who entered the registry after 2002 were used for parameter estimation. Between January 1st, 2003 and May 22nd, 2012, there were 231 patients with a mean follow-up of 1.5 years (minimum 0 year; maximum 9.2 years).

Percentages of male and female patients with paid jobs were estimated based on a Dutch cross-sectional study on labour force participation [24]. Based on these distributions, the baseline characteristics of each patient were sampled using the sampling methods described in the online supplementary appendix.

2.3.2 Treatment Effectiveness and Toxicity

Data from the Dutch RhEumatoid Arthritis Monitoring (DREAM) biologic registry [25] (BRM treatment) and the Nijmegen Inception cohort (DMARD treatment) were used to estimate: treatment effect on DAS28, time to end of DAS28 decrease and time to occurrence of severe toxicity since the start of a drug; and time to loss of response since the end of DAS28 decrease. The DREAM biologic registry was started in April 2003 to monitor and evaluate the use of BRMs in patients who had not responded to methotrexate (optimal dose) and at least another DMARD before the inclusion date from 13 hospitals in The Netherlands. Up to June 28th, 2011, 1,799 patients had been included in this cohort with a mean follow-up of 2.2 years (minimum 0 year; maximum 8.2 years).

Because of insufficient data, we assumed that the effectiveness of a specific drug was independent of the identities and the causes of failure of the drugs that had been given previously. We found a linear relationship between the absolute change in DAS28 during the decreasing phase and DAS28 at the start of a treatment. For the same BRM, we found different effectiveness between its first and second administration. Therefore, absolute changes in DAS28 were sampled for each drug or drug class, distinguishing the first and second BRM, using a statistical linear model with DAS28 at the start of the treatment as an explanatory variable (Table 2). The assumption on normal distribution of the errors was verified using the Shapiro–Wilk test [26]. Because this assumption was satisfied for any linear model for DAS28 change, the assumption on homoscedasticity of the errors was further verified using the Breusch–Pagan test [27]. We

Table 2 Model parameters for sampling patient characteristics and times to events, and estimating treatment effectiveness, and costs and QALYs

Parameters	Estimate	Distribution[a]	Source
Proportion of male patients who have paid jobs	0.85	Beta(249, 43)	Chorus et al. [24]
Proportion of female patients who have paid jobs	0.42	Beta(318, 444)	Chorus et al. [24]
Intercept (α) and slope (β) in the linear regression model for estimating absolute change in DAS28 based on DAS28 at the start of:			
Methotrexate ($n = 91$)	$\alpha = -1.08$; $\beta = 0.73$	MVN	NIC
Sulphasalazine ($n = 27$)	$\alpha = -0.98$; $\beta = 0.68$	MVN	NIC
Leflunomide ($n = 21$)	$\alpha = -1.42$; $\beta = 0.71$	MVN	NIC
Other DMARDs ($n = 18$)	$\alpha = -0.67$; $\beta = 0.32$	MVN	NIC
Adalimumab, first position in sequential use of anti-TNF BRMs ($n = 226$)	$\alpha = -0.95$; $\beta = 0.72$	MVN	DREAM
Adalimumab, second position in sequential use of anti-TNF BRMs ($n = 120$)	$\alpha = -1.48$; $\beta = 0.75$	MVN	DREAM
Etanercept or certolizumab, first position in sequential use of anti-TNF BRMs ($n = 267$)	$\alpha = -0.87$; $\beta = 0.70$	MVN	DREAM
Etanercept or certolizumab, second position in sequential use of anti-TNF BRMs ($n = 167$)	$\alpha = -0.81$; $\beta = 0.53$	MVN	DREAM
Infliximab or golimumab, first position in sequential use of anti-TNF BRMs ($n = 188$)	$\alpha = -0.33$; $\beta = 0.46$	MVN	DREAM
Infliximab or golimumab, second position in sequential use of anti-TNF BRMs ($n = 19$)	$\alpha = -0.36$; $\beta = 0.41$	MVN	DREAM
Non-TNF BRM, first position in sequential use of non-TNF BRMs ($n = 73$)	$\alpha = -0.71$; $\beta = 0.57$	MVN	DREAM
Non-TNF BRM, second position in sequential use of non-TNF BRMs ($n = 18$)	$\alpha = -0.72$; $\beta = 0.53$	MVN	DREAM
Coefficients in the linear mixed models for predicting HAQ, $HAQ = (\alpha + u_0) + \beta_1 \times DAS28 + \beta_2 \times AGE + (\beta_3 + u_1) \times DR + \beta_4 \times SEX + \varepsilon$[b] in patients receiving a:			
DMARD ($n = 185$)	$\alpha = -0.39$; $\beta_1 = 0.14$; $\beta_2 = 0.01$; $\beta_3 = 0.03$; $\beta_4 = 0.29$	MVN	NIC
BRM ($n = 1,579$)	$\alpha = -0.31$; $\beta_1 = 0.15$; $\beta_2 = 0.01$; $\beta_3 = 0.01$; $\beta_4 = 0.24$	MVN	DREAM
Scale parameter in the exponential distribution of time to end of DAS28 decrease (in year) in patients receiving a/an:			
DMARD ($n = 157$)	1.69	N(0.53, 0.09)[c]	NIC
Anti-TNF BRM ($n = 987$)	1.61	N(0.48, 0.53)[c]	DREAM
Non-TNF BRMs ($n = 91$)	0.49	N(−0.71, 0.170)[c]	DREAM
Scale parameter in the exponential distribution of time to loss of response (in year) of patients receiving a:			
DMARD ($n = 121$)	0.97	N(−0.03, 0.18)[c]	NIC
Anti-TNF BRMs ($n = 726$)	2.38	N(0.87, 0.13)[c]	DREAM
Non-TNF BRMs ($n = 52$)	0.58	N(−0.54, 0.17)[c]	DREAM
Coefficients in the parametric survival regression model for time to severe toxicity (in year), $\log(\mu) = \alpha + \beta \times AGE$, of[d]:			
Methotrexate ($n = 139$)	$\alpha = 2.46$; $\beta = -0.021$	MVN	NIC
Sulphasalazine ($n = 42$)	$\alpha = 2.95$; $\beta = -0.025$	MVN	NIC
Leflunomide ($n = 39$)	$\alpha = 1.62$; $\beta = -0.014$	MVN	NIC
Other DMARDs ($n = 23$)	$\alpha = 2.44$; $\beta = -0.019$	MVN	NIC
Anti-TNF BRMs, first position in sequential use of anti-TNF BRMs ($n = 1,136$)	$\alpha = 1.62$; $\beta = -0.012$	MVN	DREAM
Anti-TNF BRMs, second position in sequential use of anti-TNF BRMs ($n = 510$)	$\alpha = 2.24$; $\beta = -0.026$	MVN	DREAM

Table 2 continued

Parameters	Estimate	Distribution[a]	Source
Non-TNF BRMs, first position in sequential use of non-TNF BRMs ($n = 121$)	$\alpha = -0.21$; $\beta = 0.019$	MVN	DREAM
Non-TNF BRMs, second position in sequential use of non-TNF BRMs ($n = 32$)	$\alpha = -2.98$; $\beta = 0.072$	MVN	DREAM
Coefficients in the two-part model for predicting 3-month non-drug healthcare cost (€) ($n = 425$)			
1st part: logit[P(non-zero cost)] $= \alpha + \beta_1 \times$ DR $+ \beta_2 \times$ HAQ $+ \beta_3 \times$ SEX $+ \beta_4 \times$ HAQ $\times$ SEX	$\alpha = -0.87$; $\beta_1 = 0.02$; $\beta_2 = 0.78$; $\beta_3 = -0.31$; $\beta_4 = 0.67$	MVN	DREAM
2nd part: log(cost) $= \alpha + \beta_1 \times$ DAS28 $+ \beta_2 \times$ HAQ $+ \beta_3 \times$ SEX $+ \beta_4 \times$ DAS28 $\times$ SEX $+ \varepsilon$	$\alpha = 5.93$; $\beta_1 = 0.18$; $\beta_2 = 0.42$; $\beta_3 = 0.93$; $\beta_4 = -0.24$	MVN	DREAM
Coefficients in the two-part model for predicting 3-month sick leave cost (€) ($n = 425$)		MVN	DREAM
1st part: logit[P(non-zero cost)] $= \alpha + \beta_1 \times$ AGE $+ \beta_2 \times$ SEX	$\alpha = 1.90$; $\beta_1 = -0.07$; $\beta_2 = -0.91$	MVN	DREAM
2nd part: log(non-zero cost) $= \alpha + \beta_1 \times$ HAQ $+ \beta_2 \times$ SEX $+ \varepsilon$	$\alpha = 7.87$; $\beta_1 = 0.85$; $\beta_2 = -1.53$	MVN	DREAM
Coefficients in the two-part model for predicting health utility (HU) ($n = 425$)		MVN	DREAM
1st part: logit[P(class 2)] $= \alpha + \beta_1 \times$ AGE $+ \beta_2 \times$ DAS28 $+ \beta_3 \times$ HAQ	$\alpha = 4.33$; $\beta_1 = 0.04$; $\beta_2 = -0.56$; $\beta_3 = -1.69$	MVN	DREAM
2nd part: $HU_{class\ 1} = \alpha + \beta_1 \times$ HAQ $+ \beta_2 \times$ SEX $+ \varepsilon$	$\alpha = 0.17$; $\beta_1 = -0.11$; $\beta_2 = 0.04$	MVN	DREAM
$HU_{class\ 2} = \alpha + \beta_1 \times$ HAQ $+ \beta_2 \times$ SEX $+ \beta_3 \times$ DAS28 $+ \varepsilon$	$\alpha = 0.82$; $\beta_1 = -0.11$; $\beta_2 = 0.03$; $\beta_3 = -0.01$	MVN	DREAM

DMARD disease-modifying antirheumatic drug; *TNF* tumour necrosis factor; *BRM* biological response modifier; *SEX* takes on a value of 1 or 0 if the patient is female or male, respectively; *DR* disease duration; *HAQ* Health Questionnaire Assessment Disability Index; *DAS28* Disease Activity Score for 28 joints; *n* number of observations in the data set used for model fitting; *NIC* Nijmegen Inception Cohort; *DREAM* Dutch RhEumatoid Arthritis Monitoring registry

[a] Distribution used to sample parameter values in the probabilistic sensitivity analysis; $N(m, s)$, normal distribution with mean m and standard deviation s; Beta(a, b), beta distribution with scale parameters a and b, which were estimated using the method of moments [55]; MVN, multivariate normal distribution with mean and variance-covariance matrix estimated by fitting the regression models

[b] u_0 and u_1 followed bivariate normal distribution with a mean vector 0 and the variance-covariance matrix estimated by fitting the linear mixed model

[c] Distribution of the logarithm of the scale parameter; see Selvin [57] for method of parameter estimation

[d] μ is the scale parameter in the exponential distribution of time to toxicity

analysed all the DMARDs that a patient received after failure of methotrexate and sulphasalazine (or leflunomide) as one class, referred to as 'other DMARDs', because the numbers of observations for some drugs were too small ($n < 5$) and the effectiveness of the drugs with more observations were found to be similar; the effect of 'other DMARDs' on DAS28 was assumed to be the same in both treatment strategies 1 and 2. For the same reasons, all non-TNF BRMs at the same position in the treatment sequence were grouped, and specific pairs of anti-TNFs also grouped (see Table 2).

Times to end of DAS28 decrease and to loss of response followed exponential distributions. We found similar parameter values in the distributions among DMARDs, among anti-TNF BRMs and among non-TNF BRMs. Therefore, times to the above-mentioned events were sampled for different classes of drugs instead of individual drugs. Because the Pearson's correlation coefficients between the absolute changes in DAS28 and times to end of DAS28 decrease for different drugs were small (<0.3), we assumed that a change in DAS28 was independent of the time taken to achieve this change. Therefore, we sampled the time to end of DAS28 decrease separately from sampling the change in DAS28. Time to severe toxicity for each treatment was linked to age, gender and disease duration using a parametric survival model with assumed exponential distribution of the dependent variable. To obtain a parsimonious model, we performed backward variable selection using the Akaike Information Criterion (AIC) [28]. As a result, age was the only predictor of time to severe toxicity in the final model (see Table 2).

Long-term progression of physical functioning, measured by the Health Assessment Questionnaire Disability Index (HAQ) [29–31], was simulated based on the longitudinal relationship between DAS28 and HAQ. The changes in DAS28 and HAQ over time were linked to costs and

health utility (HU) using statistical models (see next sections and Table 2).

2.3.3 Costs and Quality-Adjusted Life-years

Data from the DREAM biologic registry were used for statistical modelling of healthcare and sick leave costs and HU. Costs incurred by each patient were calculated based on the amounts of resource use and unit costs. The amounts of resource use were reported every 3 months by each patient participating in the DREAM registry and included: numbers of diagnostic and laboratory tests; days in hospitals; numbers of surgical procedures; visits to rheumatologists and other specialists, to general practitioners, to nurse specialists, to physiotherapists, and to psychologists; and hours of formal and informal care. The unit costs published by the Dutch Health Care Insurance Board [32] were used and adjusted using a consumer price index [33]. The human capital approach was used for estimating sick leave cost [34].

For healthcare and sick leave costs, we used two-part models as recommended by Mullahy [35] and Manning and Mullahy [36] to link cumulative costs over 3 months to age, gender, disease duration and mean DAS28 and HAQ. Because of the excessive zeros in the cost data, logistic regression models were used in the first part to estimate the probability that the costs were non-zeros. Conditional on non-zero costs, log-linear regression models were used in the second part to evaluate the costs. The assumptions on normality and homoscedasticity of the errors in these log-linear models were verified using the same tests as in the linear models for changes in DAS28. Because the DREAM registry did not contain data related to permanent work disability (PWD), we estimated time to PWD based on Chorus et al. [24]. The productivity cost due to PWD was estimated using the human capital approach based on the duration of PWD in patients with paid jobs, and average wages derived from Statistics Netherlands [33] for men and women in different age categories. The total productivity cost was calculated as the sum of costs due to sick leave and PWD.

Total drug cost for each patient was computed based on the simulated durations that the patient received specific drugs. Doses and unit prices of the drugs were determined based on documentation of the Dutch Institute for Health Care [37].

In the DREAM registry, HU was measured using the EuroQoL five dimensions questionnaire (EQ-5D) [38, 39]. The histogram of the observed HU values showed bimodal normal distributions and therefore we performed a computer-assisted analysis of mixtures to divide the data into two latent classes [40]. Then, we used a two-part model to link HU to age, gender, disease duration, DAS28 and HAQ. In the first part, a logistic regression model was used to predict the probability that a patient belonged to a specific class. In the second part, a linear regression model, fitted for each specific class, was used to predict the HU (see Table 2). Predicted values of HU that fell outside of the realistic ranges were adjusted using the approach developed by Hernández Alava et al. [41]. In fitting each of the models for costs and HU, we performed backward variable selection using the AIC (see Table 2 for the final models).

QALYs were computed based on the predicted HU values at discrete time points. We selected QALY as an outcome because it is recommended for use as a standard measure of health effects in the Dutch and other guidelines for conducting and reporting economic evaluation [42–44]. Despite controversial assumptions underlying QALY, it has been by far the only generic measure of health that allows comparisons between diseases. Extensive discussion on the advantages and limitations of the QALY approach can be found in Loomes and McKenzie [45], Wagstaff [46], Broome [47] and Dolan [48]. Following the Dutch guidelines for pharmacoeconomic studies, annual discount rates for costs and QALYs were set at 4.0 and 1.5 %, respectively [42].

2.3.4 Predicting the Health Assessment Questionnaire Disability Index

Because costs and QALYs are non-linearly related to DAS28 and HAQ (see previous section), we predicted HAQ based on disease duration, DAS28, age and gender using a linear mixed-effects model with random effects of intercept and disease duration. According to the rheumatologists' opinion, the long-term beneficial effect on HAQ of BRMs might be larger than that of DMARDs. Therefore, the models for HAQ progression in patients receiving DMARDs and BRMs were fitted separately using the Nijmegen Inception Cohort and DREAM registry, respectively. An overview of the model parameters, their estimates and distributions for probabilistic sensitivity analysis (PSA), and data sources for parameterisation is provided in Table 2.

2.4 Modelling Tools, Model Implementation and Output Analysis

The DES model was written using the Delphi language (Embarcadero Delphi XE 15.0; Em-barcadero Technologies Inc., San Francisco, CA, USA). R [49] and SAS 9.2 (2008; SAS Institute Inc., Cary, NC, USA) were used for data handling and statistical analyses. The package C.A.MAN was used for the mixture analysis of HU [50]. The algorithms provided by Ripley [51], Genz [52] and Press et al. [53] were used to generate random variables.

The model was rigorously verified for coding logics and correctness, and debugged based on extreme value scenarios.

Because RA is a chronic disease and the sequential treatment strategies included high numbers of drugs, a lifetime horizon was chosen to capture any differences in the long-term outcomes between the treatment strategies as recommended by the ISPOR Task Force on good research practices for modelling studies [54]. Therefore, the simulation was run until death of patients. The size of the initial cohort was determined by repeatedly running the simulation with increasing initial population size until means and standard deviations (SDs) of costs and QALYs became stable. As a result, an initial cohort of 10,000 patients was used for the first- and second-order uncertainty analyses. ICER was calculated according to Briggs et al. [55]. Ninety-five-percent confidence interval (CI) of ICER was computed using the non-parametric bootstrapping method [56] with 100,000 times of sampling. Simulated results were face-to-face validated by experts in RA and health economists.

For PSA, values of the model parameters were sampled 1,000 times from appropriate distributions. The beta distributions of the proportions (see Table 2) were estimated using the method of moments [55]. The multivariate normal distributions of the parameters in the regression models were obtained from the model fitting. The distribution of the logarithm of the scale parameters in the exponential distributions were estimated based on survival analyses [57]. The net-benefit framework was used to construct the cost-acceptability curves from the Monte Carlo simulation results [55].

3 Results

3.1 Disease Progression

The simulated results showed equal or better DAS28 and HAQ (as population averages) in Strategy 2 over time compared with Strategy 1 (Fig. 2). During the first 2 years, DAS28 and HAQ in both strategies decreased rapidly. After this period, DAS28 in Strategy 1 started to increase but that in Strategy 2 continued to decrease slightly until year 6. After year 6, DAS28 in both strategies increased until year 28 and then became stable. The differences in DAS28 between the two strategies were negligible during the first 2 years, increased quickly from year 2 to year 12, decreased steadily from year 12 to year 28, and became negligible again after year 28.

Similar to the case of DAS28, HAQ (as population averages) in both strategies decreased during the first 2 years, but at a slower rate than the change in DAS28. After this period, however, HAQ in both strategies increased over time. The differences in HAQ between the two strategies were negligible during the first 2 years, increased gradually from year 2 to year 12, decreased gradually from year 12 to year 22, and became very small (<0.1) again after year 22.

3.2 Costs and Effectiveness

Mean societal costs per patient per year varied between €9,127 and €18,121 in Strategy 1, and between €10,731 and €18,027 in Strategy 2 over a period of 70 years. Overall mean (SD) of the non-drug direct cost per patient per year was €8,717 (2,937) in Strategy 1 and €8,548 (2,806) in Strategy 2, implying a marginally positive effect of BRMs on the non-drug resource use. Overall mean (SD) of the productivity costs per patient per year was €3,618 (9,101) in Strategy 1 and €3,379 (9,105) in Strategy 2. Overall mean (SD) of drug cost per patient per year was €155 (224) in Strategy 1 and €1,422 (2,242) in Strategy 2. Productivity and drug costs accounted for 48 and 4 % of the total costs in Strategies 1, respectively, and 34 and 31 % of the total costs in Strategy 2, respectively.

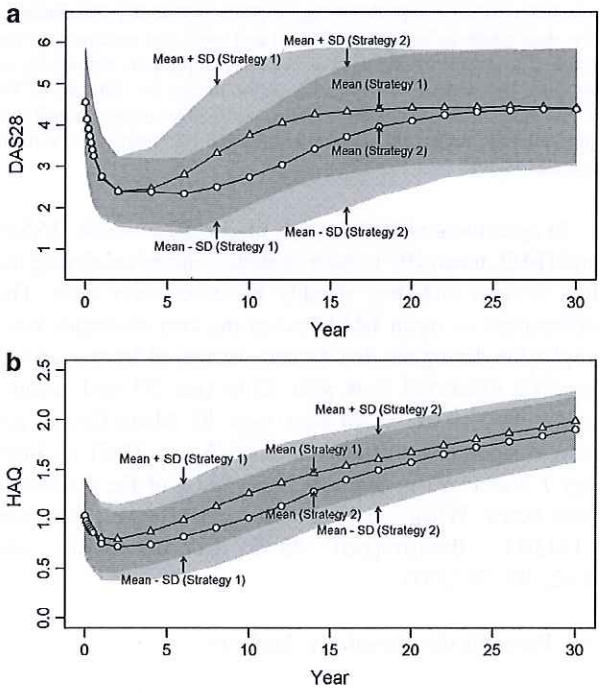

Fig. 2 Changes in means and standard deviations of DAS28 (**a**) and HAQ (**b**) of the patients in two treatment strategies over time. Strategy 1 consists of eight available DMARDs and Strategy 2 consists of the same available DMARDs as in Strategy 1 plus two anti-TNF and two non-TNF biological response modifiers. *DAS28* Disease Activity Score for 28 joints, *HAQ* Health Questionnaire Assessment Disability Index, *DMARDs* disease-modifying antirheumatic drugs, *TNF* tumour necrosis factor

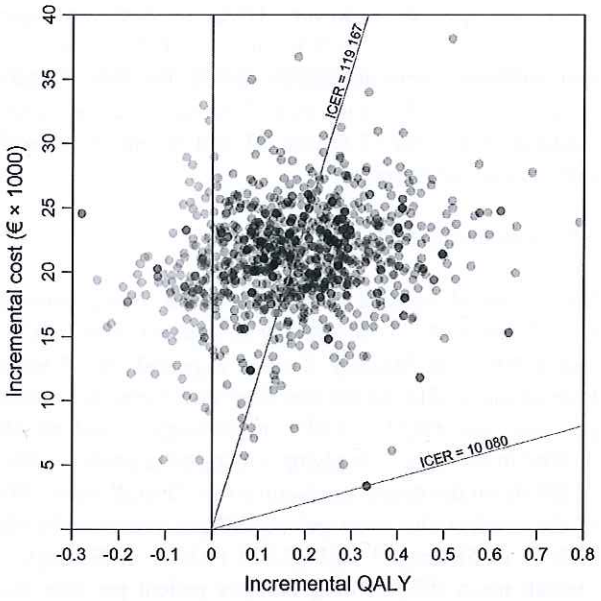

Fig. 3 Scatter plot of incremental mean costs against incremental mean QALYs of Strategy 2 compared with Strategy 1. Strategy 1 consists of eight available DMARDs and Strategy 2 consists of the same available DMARDs as in Strategy 1 plus two anti-TNF and two non-TNF biological response modifiers. Each data point was obtained from one simulation for 10,000 patients with a set of random parameter values sampled from appropriate probability distributions. The data points lie in both north-west and north-east quadrants of the plane. The *colour intensity of the dot cloud* represents the density of the dots: the stronger the intensity is, the denser the dots are. *ICER* incremental cost-effectiveness ratio, *DMARDs* disease-modifying antirheumatic drugs, *QALYs* quality-adjusted life-years, *TNF* tumour necrosis factor

In agreement with the trends of change in mean DAS28 and HAQ, mean HU in both strategies increased during the first 2 years and then steadily increased over time. The differences in mean HU between the two strategies were negligible during the first 4 years, increased from year 4 to year 12, decreased from year 12 to year 30, and became negligible (<0.01) again after year 30. Mean QALY per patient since RA diagnosis until death was 11.51 in Strategy 1 and 11.93 in Strategy 2 (95 % CI of the difference 0.08–0.86). When combing costs and effects, ICER was €124,011. Bootstrapped 95 % CI of ICER was (€62,193–597,857).

3.3 Probabilistic Sensitivity Analysis

The scatterplot of the joint uncertainty in the incremental mean costs against incremental mean QALYs of Strategy 2 compared with Strategy 1 showed that 23 % of the data points lied in the north-west and remaining points lied in the north-east quadrants of the cost-effectiveness plane (Fig. 3). At a WTP threshold of €10,080, the probability that the Strategy 2 was cost effective was zero. At a WTP

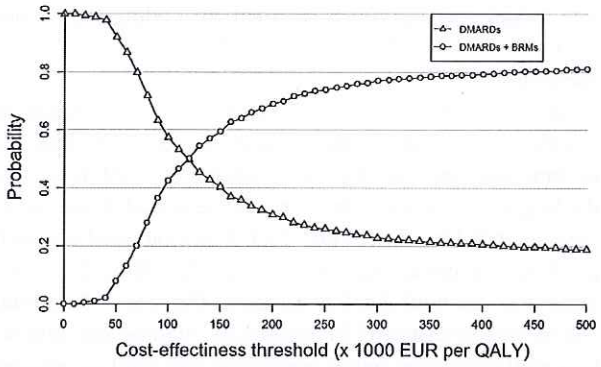

Fig. 4 Cost-effectiveness acceptability curves for two treatment strategies. Strategy 1 consists of eight available DMARDs and Strategy 2 consists of the same available DMARDs as in Strategy 1 plus two anti-TNF and two non-TNF BRMs. *DMARDs* disease-modifying antirheumatic drugs, *TNF* tumour necrosis factor, *BRMs* biological response modifiers

of €119,167, the probabilities that the Strategies 1 and 2 were cost effective were equal (0.5). At thresholds above €119,167, Strategy 2 cost-effectively dominated Strategy 1 but the probability that the Strategy 2 was cost effective never exceeded 0.87 (Fig. 4).

4 Discussion

As the treatment strategies for patients with early RA are becoming more and more complex, it is necessary to develop a modelling approach that fully supports economic evaluation of a wide range of real-life treatment strategies. While the ultimate goal of treating RA is to achieve remission, the desired treatment target may differ among countries. In our study, a clinical guideline was used to formulate the decisions when to change treatment and which drug should be used to maintain low disease activity (i.e. DAS28 $\leq$3.2) in RA patients. Low disease activity is accepted in international guidelines as a more realistic target compared with remission [58]. Using this treatment target, failure and switching of a treatment in the model were determined by DAS28, rheumatic factor and treatment duration. This approach is more realistic than using fixed sequences because it better reflects the practice of rheumatologists.

Owing to the flexible simulation approach, our model can be used to assess the effects of a number of factors on disease progression, costs and effectiveness such as: the number of BRMs available for a treatment strategy, the position of a specific drug in the sequential treatment, the cut-point of DAS28 as a criterion for switching treatment to reach low disease activity and the time interval between two assessments of disease activity. In addition, the model can be adapted to simulate the rheumatologist decisions on

drug use based on potential patient characteristics, of which values can be obtained, for example, using biomarkers. Although the model was developed for the current Dutch situation, it can be easily extended to incorporate different treatment algorithms based on new guidelines and guidelines from other healthcare systems.

In this study, we sought to estimate costs and effectiveness as they would occur in routine clinical practice. While data from controlled clinical trials are ideal for estimating treatment efficacy and for comparing outcomes between treatments in selective patients, they may not be relevant for economic evaluation studies in diverse patients under conditions different from those in the controlled trials [59]. Therefore, the use of real-life data for model inputs in our study is an advantage. Because the Nijmegen Inception Cohort and the DREAM registry were long-term follow-up studies that included patients with a wide range of disease severity, it was possible to link treatment effectiveness to DAS28 at the start of the treatment. In this way, the effectiveness is not overestimated in patients with mild RA or underestimated in patients with severe RA as in the case when the treatment effect is estimated independently of disease severity. Because the assumptions of normality and homoscedasticity of errors were satisfied for any ordinary least squares-based model fitted in our study, the predicted treatment effectiveness (using linear regression models) and costs (using log-linear regression models) were unbiased. We applied statistical models based on mixtures of parametric distributions for costs, which have been shown to perform better than those based on single distributions [60]. This approach was also applied for HU, of which the histogram clearly showed a mixture of two subpopulations. Recently, Hernández Alava et al. [41] have demonstrated that mixture models performed much better than standard linear regression models in predicting HU measured by the EQ-5D questionnaires. The simulated changes in mean DAS28 and HAQ in the two treatment strategies (Fig. 2) are plausible. During the first 2 years, most of the patients in Strategy 2 received DMARDs as those in Strategy 1. Therefore, means of these disease measures were almost identical. After 2 years, more and more patients in Strategy 2 were eligible to receive BRMs. Because the effectiveness of the first anti-TNF BRM was larger than that of DMARDs, and BRMs on average maintain DAS28 at low levels for a longer time than DMARDs, the numbers of patients having DAS28 returning to the baseline levels and of patients receiving 'palliative' therapy in Strategy 1 were increasingly higher than in Strategy 2, leading to the divergence in the differences between means of disease measures in the two strategies over the next 10 years. After 12 years, more and more patients in Strategy 2 received DMARDs after failures of all BRMs, narrowing the differences between the two strategies. After 25 years, a majority of the patients in the two strategies have lost response to all the available drugs and therefore the differences became negligible.

Our simulated ICER (€124,011) was higher than conventionally accepted WTP thresholds (e.g. £23,000 in UK and $62,000 in USA) [61]. In the models that evaluated treatment sequences independent of patient characteristics and used data from clinical trials for model parameterisation, estimated ICERs were lower than €100,000 [8, 9]. Inclusion of BRMs in a treatment strategy for RA appeared to be less favourable in our model, which used a real-life treatment algorithm and data. Analyses of the DREAM registry showed that the beneficial effect on DAS28 of BRMs was only slightly larger than those of DMARDs. Although in the long term,, a patient treated with BRMs might have better physical functioning (i.e. lower HAQ) than treated with DMARDs (given the same baseline DAS28), this difference is rather small (see statistical models for HAQ in Table 2). These statistical analyses agree well with the simulated changes in DAS28 and HAQ over time in the two treatment strategies.

It should be noted that our estimated ICER was computed based on the expected costs and utilities of the whole population. Given the fact that HAQ progression, utilities and costs of patients with different ages, gender, disease duration and DAS28 are different, net benefit of a patient receiving a specific treatment clearly depends on his or her characteristics. As a consequence, there are always patients who do not benefit from the decided treatment strategy, leading to benefit forgone from the population-based decision. Thus, making a treatment decision for individual patients or subgroups of patients with similar characteristics would optimize the population welfare. Recently, much attention has been paid to an individual-based decision and the potential value that a society is willing to pay to realize the individualized care, the so-called expected value of individualized care (EVIC) [62, 63]. As our model took patient heterogeneity into consideration, it can be used to compute EVIC.

Our modelling approach has some limitations. It is complex and thus communication of the model structure and results with the decision makers may be not straightforward. In addition, extensive data are required for model parameterisation. Very large longitudinal studies are needed to quantify the effectiveness and toxicity profile of a drug at different positions and after stopping treatment of previous drugs because of different causes (primary failure, secondary failure or severe toxicity) in a treatment sequence. Because the contemporary management of RA aims at reaching low disease activity, the use of a high number of DMARDs is adopted. However, it currently remains uncertain about the effect on DAS28 of a DMARD after failure of several other DMARDs and BRMs. Because

of insufficient data, in our model all DMARDs after failure of all BRMs were analysed as one class, and the effectiveness of the second anti-TNF and non-TNF BRMs were estimated independently of the causes of failure and of toxicity of the first ones. Additionally, time to permanent work disability was modelled independently of the types of treatment. At the present time, it is still debatable whether or not treatment with BRMs could reduce the risk of becoming permanently work disabled [64]. Nevertheless, while complexity and demand for data are limitations of our model, they are helpful in identifying the knowledge gaps in the current guidelines for the management of RA and priorities for future research.

5 Conclusions

It is possible to model the outcomes of complex treatment strategies based on a clinical guideline for RA. This modelling approach allows assessment of the impact of changes in clinical practice on disease progression, health and costs of the RA patients, and cost effectiveness of new treatment decisions. Following the Dutch guideline for the management of RA and using real-life data, the simulated ICER of the treatment strategy including BMRs compared with a treatment strategy without BMRs was higher than conventionally accepted WTP thresholds. Using an individual-oriented approach, our model can serve as a tool for individualized care analysis taking patient heterogeneity into account and can be extended to assess total health and budget impact of treatment strategies in different decision-making contexts. Despite complexity and demand for extensive data, our modelling approach can help to identify the knowledge gaps in clinical guidelines for RA management and priorities for future research.

Financial disclosure/conflict of interest ATD, WK and PLCMvR declared no conflicts of interest. AB received research grants from Merck, AbbVie and Amgen, and honorarium from Pfizer and VCB. MAFJvdL received honorarium from Abbott, AbbVie, Bristol-Myers Squibb, MSD and Pfizer, fees for participation in review activities from Abbott, AbbVie, Bristol-Myers Squibb, Pfizer and UCB, and payment for lectures from Bristol-Myers Squibb. JLS received a research grant from Pfizer for conducting this study.

Role of the sponsors Pfizer had no role in development of the model, interpretation of the simulated results or writing of the manuscript.

Author contributions ATD designed and programmed the model, reviewed literature and performed statistical analyses for model inputs, ran the simulations, debugged the model, analysed the model outputs, designed and created the figures, and drafted the manuscript. AB assisted in drafting the manuscript. WK compiled the input data and estimated costs and utility. AB, WK, PLCMvR and MAFJvdL provided the first algorithm for treatment decisions. JLS assisted in conceptualising the model. All authors participated in discussion and refinement of the algorithm for treatment decisions, interpretation of the simulated results and review of the manuscript. ATD will serve as a guarantor for the entire contents of the manuscript.

References

1. Pincus T, Yazici Y, Sokka T, Aletaha D, Smolen JS. Methotrexate as the "anchor drug" for the treatment of early rheumatoid arthritis. Clin Exp Rheumatol. 2003;21(5 Suppl 31):179–85.
2. Saag KG, Teng GG, Patkar NM, Anuntiyo J, Finney C, Curtis JR, et al. American College of Rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. Arthritis Care Res. 2008; 59(6):762–84.
3. Smolen JS, Landewé R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis. 2010; 69(6):964–75.
4. Emery P, Van Vollenhoven R, Ostergaard M, Choy E, Combe B, Graninger W, et al. Guidelines for initiation of anti-tumour necrosis factor therapy in rheumatoid arthritis: similarities and differences across Europe. Ann Rheum Dis. 2009;68(4):456–9.
5. Weinstein MC, Torrance G, McGuire A. QALYs: the basics. Value Health. 2009;12(Suppl 1):S5–9.
6. Prevoo MLL, Van't Hof MA, Kuper HH, Van Leeuwen MA, Van de Putte LBA, Van Riel P. Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38(1):44–8.
7. Fransen J, van Riel P. The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol. 2005;23(5):93.
8. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technol Assess. 2004;8:1–91.
9. Davies A, Cifaldi MA, Segurado OG, Weisman MH. Cost-effectiveness of sequential therapy with tumor necrosis factor antagonists in early rheumatoid arthritis. J Rheumatol. 2009;36(1):16–26.
10. Saraux A, Gossec L, Goupille P, Bregman B, Boccard E, Dupont D, et al. Cost-effectiveness modelling of biological treatment sequences in moderate to severe rheumatoid arthritis in France. Rheumatology (Oxford). 2010;49(4):733–40.
11. Puolakka K, Blafield H, Kauppi M, Luosujarvi R, Peltomaa R, Leikola-Pelho T, et al. Cost-effectiveness modelling of sequential biologic strategies for the treatment of moderate to severe rheumatoid arthritis in Finland. Open Rheumatol J. 2012;6:38–43.
12. Brennan A, Bansback N, Nixon R, Madan J, Harrison M, Watson K, et al. Modelling the cost effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. Rheumatology (Oxford). 2007;46(8):1345–54.
13. Cimmino MA, Leardini G, Salaffi F, Intorcia M, Bellatreccia A, Dupont D, et al. Assessing the cost-effectiveness of biologic agents for the management of moderate-to-severe rheumatoid arthritis in anti-TNF inadequate responders in Italy: a modelling approach. Clin Exp Rheumatol. 2010;29(4):633–41.
14. Beresniak A, Ariza-Ariza R, Garcia-Llorente JF, Ramirez-Arellano A, Dupont D. Modelling cost-effectiveness of biologic treatments based on disease activity scores for the management of rheumatoid arthritis in Spain. Int J Inflam. 2011;2011:727634.
15. Russell A, Beresniak A, Bessette L, Haraoui B, Rahman P, Thorne C, et al. Cost-effectiveness modeling of abatacept versus other biologic agents in DMARDS and anti-TNF inadequate

responders for the management of moderate to severe rheumatoid arthritis. Clin Rheumatol. 2009;28(4):403–12.
16. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Econ. 2003;12:837–48.
17. Caro JJ. Pharmacoeconomic analyses using discrete event simulation. PharmacoEconomics. 2005;23:323–32.
18. Stahl J. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. PharmacoEconomics. 2008;26:131–48.
19. Tran-Duy A, Boonen A, van de Laar MAFJ, Franke AC, Severens JL. A discrete event modelling framework for simulation of long-term outcomes of sequential treatment strategies for ankylosing spondylitis. Ann Rheum Dis. 2011;70:2111–8.
20. Welsing PMJ, van Gestel AM, Swinkels HL, Kiemeney LALM, Van Riel PLCM. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum. 2001;44(9):2009–17.
21. van Gestel AM, Haagsma CJ, van Riel PLCM. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum. 2004;41(10):1845–50.
22. Welsing PMJ, van Riel PLCM. The Nijmegen inception cohort of early rheumatoid arthritis. J Rheumatol. 2004;69:14–21.
23. Welsing PMJ, Fransen J, van Riel PLCM. Is the disease course of rheumatoid arthritis becoming milder? Time trends since 1985 in an inception cohort of early rheumatoid arthritis. Arthritis Rheum. 2005;52(9):2616–24.
24. Chorus AMJ, Miedema HS, Wevers CJ, van der Linden S. Labour force participation among patients with rheumatoid arthritis. Ann Rheum Dis. 2000;59(7):549–54.
25. DREAM. Dutch RhEumatoid Arthritis Monitoring. 2014. http://www.dreamregistry.nl. Accessed 15 Jan 2014.
26. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika. 1965;52(3/4):591–611.
27. Breusch TS, Pagan AR. A simple test for heteroscedasticity and random coefficient variation. Econometrica. 1979;47(5):1287–94.
28. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov EBN, Caski F, editors. Second international symposium on information theory. Budapest: Akademiai Kiado; 1973. p. 267–281.
29. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980;23(2):137–45.
30. Ramey DR, Raynauld JP, Fries JF. The health assessment questionnaire 1992: status and review. Arthritis Rheum. 1992;5(3):119–29.
31. Bruce B, Fries JF. The health assessment questionnaire (HAQ). Clin Exp Rheumatol. 2005;23(5):14.
32. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding voor kostenonderzoek: methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. Diemen: College voor Zorgverzekeringen; 2010.
33. CBS. Statline-Centraal Bureau voor de Statistiek. 2014. http://statline.cbs.nl/statweb/. Accessed 1 Feb 2014.
34. Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. PharmacoEconomics. 1996;10(5):460–6.
35. Mullahy J. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. J Health Econ. 1998;17(3):247–81.
36. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? J Health Econ. 2001;20(4):461–94.
37. Zorginstituut Nederland. Farmacotherapeutisch Kompas 2014. http://www.farmacotherapeutischkompas.nl. Accessed 1 Feb 2014.
38. The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy. 1990;16:199–208.
39. Dolan PD, Gudex C, Kind P, Williams A. A Social tariff for EuroQoL: results from a UK general population survey. Discussion Paper No. 138. York: Centre for Health Economics, University of York; 1995.
40. Böhning D, Schlattmann P, Lindsay B. Computer-assisted analysis of mixtures (C.A.MAN): statistical algorithms. Biometrics. 1992;48(1):283–303.
41. Hernández Alava M, Wailoo AJ, Ara R. Tails from the peak district: adjusted limited dependent variable mixture models of EQ-5D questionnaire health state utility values. Value Health. 2012;15(3):550–61.
42. College voor Zorgverzekeringen. Richtlijnen farmaco-economisch onderzoek, geactualiseerde. Versie 2006.
43. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the panel on cost-effectiveness in health and medicine. JAMA. 1996;276(15):1253–8.
44. National Institute for Health and Clinical Excellence. Guide to the methods for technology appraisal. London: NICE; 2008. http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf. Accessed 10 June 2014.
45. Loomes G, McKenzie L. The use of QALYs in health care decision making. Soc Sci Med. 1989;28(4):299–308.
46. Wagstaff A. QALYs and the equity-efficiency trade-off. J Health Econ. 1991;10(1):21–41.
47. Broome J. QALYs. J Public Econ. 1993;50(2):149–67.
48. Dolan P. The measurement of individual utility and social welfare. J Health Econ. 1998;17(1):39–52.
49. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-02012.
50. Böhning D, Dietz E, Schlattmann P. Recent developments in computer-assisted analysis of mixtures. Biometrics. 1998;54(2):525–36.
51. Ripley BD. Stochastic simulation. New York: Wiley; 1987.
52. Genz A. Numerical computation of multivariate normal probabilities. J Comput Graph Stat. 1992;1(2):141–9.
53. Press WH, Teukolsky SA, Vetterling W, Flannery BP. Numerical recipes-the art of scientific computing. 3rd ed. New York: Cambridge University Press; 2007.
54. 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 Health. 2003;6(1):9–17.
55. Briggs A, Claxton K, Sculpher M. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
56. Briggs A, Wonderling D, Mooney C. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ. 1997;6:327–40.
57. Selvin S. Survival analysis for epidemiologic and medical research: a practical guide. Cambridge: Cambridge University Press; 2008.
58. Smolen JS, Aletaha D, Bijlsma JWJ, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis. 2010;69(4):631–7.
59. Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in economic evaluation: an unavoidable fact of life. Health Econ. 1997;6(3):217–27.
60. Mihaylova B, Briggs A, O'Hagan A, Thompson SG. Review of statistical methods for analysing healthcare resources and costs. Health Econ. 2011;20(8):897–916.
61. Shiroiwa T, Sung YK, Fukuda T, Lang HC, Bae SC, Tsutani K. International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? Health Econ. 2010;19(4):422–37.

62. Basu A, Meltzer D. Value of information on preference heterogeneity and individualized care. Med Decis Mak. 2007;27(2):112–27.
63. van Gestel A, Grutters J, Schouten J, Webers C, Beckers H, Joore M, et al. The role of the expected value of individualized care in cost-effectiveness analyses and decision making. Value Health. 2012;15(1):13–21.
64. Ter Wee MM, Lems WF, Usan H, Gulpen A, Boonen A. The effect of biological agents on work participation in rheumatoid arthritis patients: a systematic review. Ann Rheum Dis. 2012;71(2):161–71.