PharmacoEconomics (2014) 32:429–442
DOI 10.1007/s40273-014-0139-9

SYSTEMATIC REVIEW

How to Select the Right Cost-Effectiveness Model?

A Systematic Review and Stepwise Approach for Selecting a Transferable Health Economic Evaluation Model for Rheumatoid Arthritis

H. G. M. van Haalen · J. L. Severens · A. Tran-Duy · A. Boonen

Published online: 7 February 2014
© Springer International Publishing Switzerland 2014

Abstract

Objective In the current study, we propose an approach for selection of a model that is transferable to a specific decision-making context (in this case, the Netherlands), using the case of rheumatoid arthritis (RA). The objectives of this study were (a) to perform a systematic literature review to identify existing health economic evaluation models for economic evaluation of disease-modifying antirheumatic drugs (DMARDs) in RA; and (b) to test the appropriateness of a stepwise model-selection process.

Methods First, we searched Medline and Embase to identify relevant studies in the English language, published between 1 January 2002 and 31 August 2012. From the included studies, all unique models were identified. Second, we applied a multi-step approach to model selection. Models that did not meet all minimal methodological and structural requirements based on the Outcome Measures in Rheumatology (OMERACT) criteria were excluded. Next, models were assessed on the basis of their fit when transferred to the Dutch health care setting. The criteria for model fit were transferability factors, as published by Welte et al., after exclusion of those that were deemed transferable by simple adaptation. Finally, the remaining models underwent a general quality check using the Philips checklist. Models showing good fit and high quality were considered to be transferable to the Dutch health care setting, using simple adaptation.

Results The systematic literature search identified 498 articles, which included 33 unique health economic evaluation models. Only six models passed the minimal methodological and structural requirements. Two of these models had an imperfect transferability fit to the Dutch health care setting, according to the Welte method. The remaining four models were, according to the Philips method, of good quality and were expected to be transferable by a simple adaptation.

Conclusion This study introduces a stepwise approach for selecting health economic evaluation models that are transferable by a simple adaptation. The approach seems feasible and can be applied in various therapeutic areas, provided that the minimal methodological and structural requirements are defined accordingly. Availability of health economic evaluation models coupled with structured model selection could improve the efficiency, quality and comparability of health economic research.

Electronic supplementary material The online version of this article (doi:10.1007/s40273-014-0139-9) contains supplementary material, which is available to authorized users.

H. G. M. van Haalen (✉) · J. L. Severens
Institute of Health Policy and Management, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands
e-mail: vanhaalen@bmg.eur.nl

H. G. M. van Haalen
AstraZeneca, Zoetermeer, The Netherlands

A. Tran-Duy · A. Boonen
Division of Rheumatology, Department of Internal Medicine, Maastricht University Medical Center and Caphri School for Public Health and Primary Care, P. Debyelaan 25, 6229 HX Maastricht, The Netherlands

A. Tran-Duy
Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Center, Maastricht, The Netherlands

Key Points for Decision Makers

In health economic research, it is common practice that health economic evaluation models are newly developed to study a particular decision problem, while suitable models may already exist from previous research.

Prior to developing a new model, a structured approach to identify existing models, and a stepwise selection procedure to determine whether these models could potentially be used to study the new decision problem, could be applied.

Transparency in modelling methods and availability of health economic evaluation models could lead to more efficient, verifiable, comparable and qualitative health economic research.

1 Introduction

Keeping health care affordable is increasingly challenging in many countries. Much attention is therefore devoted to evaluating the balance between costs and effects of health care innovation. Nowadays, the majority of such economic evaluations are performed using health economic evaluation models, in which sources of clinical evidence, information on resource use, health care costs and quality of life are synthesized in order to calculate the incremental cost and effects associated with a certain medical technology in comparison with an alternative. By comparing multiple alternatives, important information on the cost effectiveness of treatments can be obtained. However, development of such health economic evaluation models requires a great amount of time and (technical) expertise.

In situations where health economic modelling is required to support decision making, health economists may face the choice to either develop a new model or reuse an existing model. Although the latter option sounds efficient, health economic evaluation models are usually not directly generalizable to other jurisdictions, because structural aspects, as well as the parameters of a model, are often jurisdiction specific [1]. It may, however, be possible to transfer an existing model, i.e. to use local data for estimation of model parameters, thereby customizing the model to a local setting on the basis of a simple adaptation [2]. The International Society For Pharmacoeconomics and Outcomes Research (ISPOR) Task Force on the Transferability of Economic Evaluations Across Jurisdictions has suggested a stepwise approach to support health economists to determine whether simple or complex adaptation of a health economic evaluation model to the jurisdiction of interest is needed. In this approach, the potential models are checked for their relevance to the decision problem, methodological soundness and comparability of the treatment patterns between the jurisdictions [2]. In addition to this approach, several checklists have been suggested in the literature to assess the transferability of specific health economic evaluation results to other jurisdictions [3]. One of these checklists is the decision chart of Welte [4], which uses knock-out criteria to determine whether transferring a model is impossible or too difficult, and therefore whether conducting a new study is a better option. Three general knock-out criteria relate to the decision problem, the comparator and the study quality, and 14 specific knock-out criteria consider other aspects such as perspective, discount rate, prices and practice variation.

While there are a considerable number of studies focusing on assessing the transferability of a specific model, little effort has been made to develop methods for selecting the most suitable model among multiple potentially transferable models, i.e. to select the model that needs the simplest adaptation in order to be transferred to another jurisdiction. For example, one may be interested in performing a health economic evaluation of disease-modifying antirheumatic drugs (DMARDs) for the treatment of rheumatoid arthritis (RA). As many health economic evaluation models have been developed for this purpose during the last 10 years—and as valid models are highly important, given the increasing complexity of treatment strategies and the advent of several new but highly expensive drugs—it could be worthwhile to investigate the transferability potential of these models and to select a model that is most suitable for simple adaptation to inform decisions in a specific jurisdiction.

In the current study, we propose an approach for the selection of a model that is transferable to a specific decision-making context (in this case, the Netherlands), using the case of synthetic and or biological disease-modifying drugs in RA (see Box 1). The objectives of this study were (a) to perform a systematic literature review to identify existing health economic evaluation models for the economic evaluation of DMARDs in RA; and (b) to test the appropriateness of a stepwise model-selection process.

Box 1

Rheumatoid Arthritis

Rheumatoid arthritis (RA) is an inflammatory joint disease, which is induced by immune intolerance. If untreated, the resulting chronic synovial inflammation causes destruction of the cartilage and erosions of the articular bony surfaces, leading to joint deformity and destruction. In Northern European countries, approximately 41 in every 100,000 people are diagnosed with RA each year, the majority of whom are female [5, 6]. Although a curative treatment is still not available for RA, a wide range of pharmacological therapies has become available in recent decades, which can suppress inflammation and even lead to remission, and can consequently prevent further loss of joint tissues and functional decline of the patients. The introduction of drugs that have all of these properties—the so-called disease modifying antirheumatic drugs (DMARDs)—has greatly improved the quality of life of people with RA, especially the biological DMARDs (bDMARDs) such as tumour necrosis factor (TNF)-a inhibitors (i.a. etanercept, infliximab, adalimumab), abatacept and rituximab [7]. However, the introduction of bDMARDs has also resulted in a steep increase of treatment costs for RA, as bDMARDs have a serious price tag [8]. This indicates the relevance of health economic evaluation in RA, as governments and other health care payers increasingly ask for justification of these prices in terms of added clinical value.

Treatment of RA in the Netherlands

In the Netherlands, patients with RA receive specialized care by rheumatologists. In order to prevent joint damage and achieve clinical remission, treatment needs to be timely and effective, striving for low disease activity. Next to physiotherapy and exercise therapy, patients start synthetic DMARD (sDMARD) therapy (the first-line treatment is usually methotrexate) 6–12 weeks after disease onset. According to the usual 'step-up' strategy, a combination of methotrexate and a bDMARD should be considered only when a combination of two sDMARDs remains ineffective. Adalimumab, infliximab or etanercept are preferred as a second-line treatment, although tocilizumab is also used as an initial bDMARD. Beyond second-line treatment, different methotrexate/bDMARD combinations (including abatacept and rituximab) should be considered. Disease activity and progression are strictly monitored during frequent outpatient visits (± every 3 months), and treatment regimens are adjusted if necessary. In general, bDMARDs are discontinued if there is no clinical response after 3 months of treatment.

2 Methods

2.1 Published Models

To identify published models for the economic evaluation of DMARDs for the treatment of RA, we systematically searched the Medline and Embase databases, as well as the National Health Service Economic Evaluation Database (NHS EED). Studies on modelling methodology and cost effectiveness published in the English language between 1 January 2002 and 31 August 2012 were identified using a range of search terms (see Appendix 1 in the Electronic Supplementary Material). Our Medline search, in a shortened version, was as follows: ((rheumat* OR reumat*) AND (arthrit* OR artrit*) AND (cost* OR economic* OR pharmacoeconomic* OR (cost AND (analys* OR benefit OR effectiveness OR utility))) AND (simulation OR model* OR (decision AND (analys* OR analytic))) AND (DMARD* OR antirheumatic* OR antireumatic* OR biologic* OR TNF OR necrosis factor* OR *alpha OR *alfa OR gold OR auranofin OR methotrexate OR MTX OR cyclosporin OR ciclosporin OR *penicillamine OR leflunomide OR azathioprine OR sulfasalazine OR SSZ OR *chloroquine OR minocycline OR etanercept OR infliximab OR adalimumab OR golimumab OR certolizumab OR anakinra OR tocilizumab OR abatacept OR rituximab OR tofacitinib)). The Embase search was conducted using the same search terms as those used in Medline. In the NHS EED, we used the broad search terms 'rheumatoid' AND 'arthritis' to ensure high sensitivity of the search. References of retrieved papers and other systematic reviews were also checked to identify additional studies. Search results were exported to Reference Manager® Professional Edition Version 12 software (Thomson Reuters, New York, NY, USA), and duplicate articles were removed.

The title and abstract of all retrieved studies were independently scanned by two reviewers (HvH and JS or ATD) on the basis of the following exclusion criteria: (1) no primary data (e.g. review papers, editorials, comments); (2) no adult study population with RA; (3) no study population starting/on DMARDs; (4) no use of a model for cost-effectiveness evaluation of DMARD therapy; (5) no examination of both cost and consequences; and (6) no associated published full text. These exclusion criteria were not applied to methodological papers on the development of a health economic evaluation model for DMARDs in RA, i.e. these papers were included regardless of whether or not they reported a specific patient population or simulation results. In cases of disagreement between reviewers, both reviewers read and discussed the full text in order to reach a consensus.

After the title and abstract scan, the full texts of the included papers were reviewed by two independent authors (HvH and JS or ATD) using the same exclusion and inclusion criteria as mentioned above. In order to select a series of unique models, multiple papers based on the same model were clustered as one case. The study selection process, including assessment of exclusion and inclusion criteria, was systematically recorded in a spreadsheet and presented by a flow chart.

2.2 Selecting a Transferable Model

2.2.1 Step 1: Which Models Are Suitable?

As the first step in determining which models are potentially suitable to be transferred, we assessed the conceptual validity of the identified models, i.e. whether the individual models adequately represented the concept of the disease and its clinical context in their modelling framework. For this purpose, it was necessary to specify a list of disease-specific minimal methodological and structural requirements that were considered important for securing the conceptual validity of health economic evaluation models.

In the case of RA, such requirements were previously specified by the Outcome Measures in Rheumatology (OMERACT) Group, which is an international organization of clinicians, researchers, methodologists, patients and industrial partners aiming at standardizing outcomes, including economic outcomes in rheumatologic diseases [9]. In 2003, the OMERACT Group published a consensus-based reference case on health economic modelling methods in RA, in which several methodological topics were discussed and guidance was provided on the preferred modelling methodology for each topic if a consensus was reached [9]. This OMERACT reference case served as a basis for our requirements. We selected those topics from the reference case that concerned conceptual validity of health economic evaluation models and for which a clear guidance was provided.

Additional methodological guidance on modelling methods in RA was provided in later publications by Bansback et al. [10, 11]. This guidance was considered to be supplemental to the OMERACT guidance and was therefore also adopted in our requirements. The final set of minimal methodological and structural requirements that was applied is summarized in Table 1. According to the OMERACT criteria, authors should clearly define the study population for which the model is developed, because clinical inputs, as well as the model structure, are often specific to a certain RA subpopulation, such as patients who do not respond to a synthetic DMARD (sDMARD) or a tumour necrosis factor (TNF)-α inhibitor [9]. Secondly, realistic treatment sequences should be modelled, as RA treatment in clinical practice includes a sequence of agents rather than a single agent [9]. Although OMERACT does not state this explicitly, the authors agreed that the sequence needed to consist of consecutive specific treatments rather than applying a subsequent treatment line consisting of several pooled treatments, as the clinical response to various treatment options could be influenced by previous therapy lines, and unique treatments are more adjustable. Third, as advised by Bansback et al. [11], withdrawal of therapy within the model should be due to lack of efficacy or severe toxicity, rather than keeping all subjects on treatment until the end of the simulation or until a specific point in time during the simulation. Such withdrawal schemes do not reflect clinical practice and could therefore jeopardize the validity of the simulation.

Next, the models should meet the following criteria: the time horizon is at least 1 year, with the possibility to extrapolate beyond this period; mortality is accounted for in the simulation [9]; and QALYs are used as a measure of

Table 1 Minimal methodological and structural requirements for securing conceptual validity

Requirements	Based on
1. A clear definition of the underlying population should be included	OMERACT [9]
2. Realistic (and explicit) treatment sequences should be modelled, depending on the clinical setting and the baseline characteristics of the patients	OMERACT [9, 11] and Bansback [9, 11]
3. Withdrawal due to lack of efficacy or severe toxicity should be modelled (i.e. a patient should not remain on a specific drug until the end of the simulation)	Bansback [11]
4. Models should include at least a 1-year time horizon, with the possibility to extrapolate beyond this period	OMERACT [9]
5. The risk of mortality should be included	OMERACT [9]
6. QALYs should be reported in combination with disease-specific outcome measures (e.g. ACR20, improvement in DAS or HAQ score)	OMERACT [9]
7. Clinical response should be defined using absolute health outcomes (DAS or HAQ score) rather than relative health outcomes (ACR response)	Bansback [10]

ACR American College of Rheumatology, *DAS* Disease Activity Score, *HAQ* Health Assessment Questionnaire, *OMERACT* Outcome Measures in Rheumatology, *QALY* quality-adjusted life-year

effectiveness to enable comparisons of cost-effectiveness outcomes across therapeutic areas [9].

Finally, it was required that progression through the model is induced by absolute changes in the subject's disease status—using, for example, the Health Assessment Questionnaire (HAQ) score or the Disease Activity Score (DAS)[1] [10]. On this line, American College of Rheumatology (ACR) responses are considered unsuitable to define transition probabilities, as in clinical practice, (dis)continuation decisions are unlikely to be informed by ACR response and such models are therefore unlikely to reflect reality [10]. Also subjects' disease status within the model, which informs estimates of resource utilization (costs) and utilities, should not be defined by ACR response status, as this reflects a relative improvement rather than the actual disease status of a patient [10].

Positive assessment based on full information on each criterion was required for selection of a model. Thus, if a model did not meet all criteria, it was excluded. Model selection was performed by one reviewer (HvH), and exclusion of models was verified by a second reviewer (JS).

2.2.2 Step 2: Which Models Fit?

Next, we specified a list of criteria to assess the model fit, thereby evaluating whether the model structure is appropriate in the context of a particular health care setting. The Dutch health care setting was used as an illustrative case. These criteria were based on the 'specific knock-out criteria' proposed by Welte et al. [4]. We distinguished between specific knock-out criteria that concerned parameter or structural uncertainty within a model. Parameter uncertainty relates to the uncertainty surrounding the input data and probabilities that govern the outcomes because of, for example, multiple (conflicting) studies, lack of internal or external validity of the study data, or lack of data [12]. Structural uncertainty is present when there is uncertainty about the functional form of the model, i.e. whether the model adequately reflects reality surrounding the decision problem. Generally, in the presence of structural uncertainty, one cannot be certain that the produced results are valid, even if the true values of all input parameters are known [12].

We defined performing a 'simple' adaptation of a health economic evaluation model as minimizing the parameter uncertainty by adapting those input values that are subject to bias when transferring the model. In this regard, when selecting a model, there seems to be no need for exclusion of a model on the basis of parameter uncertainty, as this can be addressed during the simple adaptation process. In general, it is more difficult to adapt structural uncertainty, as this often requires technical amendments to the model. Therefore, selecting a model with limited structural uncertainty would facilitate a simple model adaptation.

Table 2 Categorization of Welte's transferability factors

Transferability factors	Parameter uncertainty	Structural uncertainty
Perspective	✔	–
Discount rate	✔	–
Medical cost approach	✔	–
Productivity cost approach	✔	–
Absolute and relative prices	✔	–
Practice variation	✔	✔
Technology availability	✔	✔
Disease incidence/prevalence	✔	–
Case mix	✔	✔
Life expectancy	✔	–
Health status preferences	✔	–
Acceptance, compliance and incentives for patients	✔	–
Productivity and absenteeism	✔	–
Disease spread	✔	✔

✔ uncertainty applicable, – uncertainty not applicable

On the basis of this theory, we assessed which of Welte's specific knock-out criteria related to parameter uncertainty and which related to structural uncertainty within a model. The result of this assessment is presented in Table 2 (see Appendix 2 in the Electronic Supplementary Material for the rationale). Next, studies were excluded merely on the basis of the criteria that concerned structural uncertainty.

2.2.3 Step 3: Which Model Is Best?

In order to know which of the eligible models has the highest quality, models that passed step 2 were additionally assessed on general methodological quality, on the basis of the Philips checklist [13]. Using the Philips checklist, the models are, among other criteria, evaluated on the substantiation of the applied methodology, the model functionality, the mathematical logic and the external validity. To our knowledge, the Philips checklist is a very suitable general quality checklist for health economic modelling studies, although minor overlap between the Philips criteria and the minimal methodological and structural requirements is present (e.g. criteria S7 and S8). As the Philips criteria can be applied to all health economic evaluation studies regardless of disease area, the criteria are

[1] Some models use survival statistics based on observational data to predict the subjects' time on treatment. We also considered this a suitable approach for modelling health state transitions (i.e. treatment discontinuation).

formulated in general terms—for example, 'is the time horizon of the model sufficient (…)?' [13]. What time horizon is sufficient in the case of RA has been defined by OMERACT and is reflected in the minimal methodological and structural requirements. Thus, models that have passed the minimal methodological and structural requirements will, by definition, pass the overlapping Philips criteria as well.

3 Results

3.1 Published Models

The systematic literature search identified 498 articles, of which 55 articles were initially selected for full-text review after the title and abstract scan (Fig. 1). After review of the full texts, eight articles were additionally excluded either because they appeared to be abstracts (2) or duplicates (3), or because no model was used (2). One full paper could not be retrieved [14]. Two papers were added to the list as a result of additional reference screening. This resulted in 49 included papers, from which 33 unique health economic evaluation models were identified (as some papers concerned the same model). The defined model clusters are presented in Table 3.

3.2 Selecting a Transferable Model

3.2.1 Step 1: Which Models Are Suitable?

Out of the 33 unique models, only six models passed the minimal methodological and structural requirements [17, 24, 29, 38, 39, 54]. Ten models did not pass the requirement on treatment sequence and/or withdrawal; these models assumed either treatment discontinuation at a fixed point in time [32, 62], no subsequent treatment [25, 26], switching to a single DMARD therapy for the remainder of the time horizon after treatment discontinuation (i.e. not a realistic treatment sequence) [42, 44, 55, 58] or a subsequent therapy line consisting of 'pooled' therapies, which were non-explicit [32, 50]. Three models did not include the possibility to extrapolate beyond a time horizon of 1 year [25, 26, 37], whereas 12 models did not simulate the mortality of the patients [16, 25, 26, 41, 45–48, 51, 53, 59]. Six models did not report quality-adjusted life-years (QALYs) as an outcome [16, 26, 45–47, 51] and, finally, 15 models did not use absolute health outcomes to define health states or to define transitions through the model [15, 23, 25–28, 30, 31, 41, 43, 44, 48, 49, 52, 53]. Thirteen models were excluded because they did not meet multiple requirements, leaving six models with adequate conceptual validity to pass through to the next step.

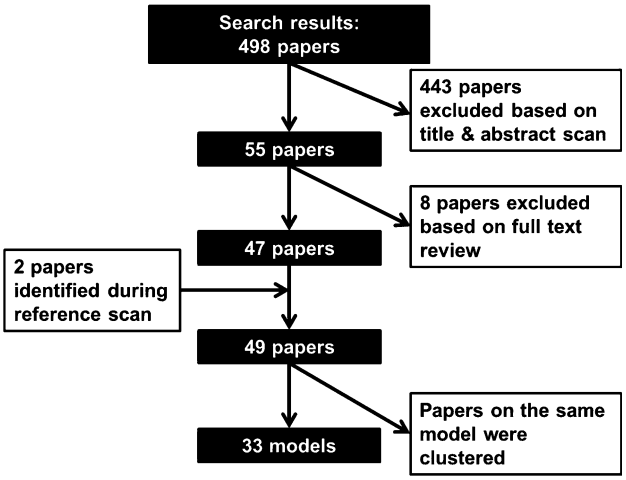

Fig. 1 Flow diagram of search results

3.2.2 Step 2: Which Models Fit?

The models by Kobelt et al. in 2011 [38] and Finckh et al. [29] were excluded on the basis of the specific knock-out criteria. Potential case-mix bias (due to differences in patient populations and therefore possible differences in treatment responses [4]) can be avoided when transferring these models through adaptation of the parameter inputs. However, the cycle length of 6 months used in both models could introduce practice variation bias when transferring them because in the Netherlands, the follow-up interval is generally 3 months, and so a cycle length of 3 months would have been more appropriate [29, 38].

Brennan et al. [24] used observational data from the British Society for Rheumatology Biologics Registry (BSRBR) to construct regression models for prediction of the European League Against Rheumatism (EULAR) response (a measure of treatment effect), impact on health utility, time on treatment and hospital stay on the basis of various patient and disease characteristics. Lindgren et al. [39] and Kobelt et al. [40] used a similar approach in their models (referred to as the 'Lindgren cluster'), in which they used data from the Southern Swedish Arthritis Treatment Group (SSATG) Register to predict time on treatment, time to treatment re-initiation and HAQ progression. Although the model predictions can be adjusted for important covariates, bias could be introduced in the transferred model if the population and clinical practice characteristics underlying the models for UK and Sweden differ from those for the Netherlands. For those simulated results that are deemed not comparable between the jurisdictions, new data analysis and adjustment of the prediction models, using Dutch-specific observational data, are necessary to avoid uncertainty due to case-mix and practice

Table 3 Results of applying the general knock-out criteria

Author/model cluster	Year	Reference	Country	Minimal methodological and structural requirements						
				Population	Sequence	Withdrawal	Time horizon	Mortality	QALYs	Absolute outcomes
Bansback	2005	[15]	Sweden	✔	✔	✔	✔	✔	✔	–
Beresniak	2011	[16]	Spain	✔	✔	✔	✔	–	–	✔
BRAM cluster	2002–2011	[17–22]	UK	✔	✔	✔	✔	✔	✔	✔
Brennan	2004	[23]	UK	✔	✔	✔	✔	✔	✔	–
Brennan	2007	[24]	UK	✔	✔	✔	✔	✔	✔	✔
Chiou	2004	[25]	US	–	–	–	–	–	✔	–
Choi	2002	[26]	US	✔	–	–	–	–	–	–
Davies	2009	[27]	US	✔	✔	✔	✔	✔	✔	–
Diamantopoulos	2012	[28]	Italy	✔	✔	✔	✔	–	✔	–
Finckh	2009	[29]	US	✔	✔	✔	✔	✔	✔	✔
Hallinen	2010	[30]	Finland	✔	✔	✔	✔	✔	✔	–
Kielhorn	2008	[31]	UK	✔	✔	✔	✔	✔	✔	–
Kobelt cluster	2002–2010	[32–37]	Sweden and UK	✔	–	–	✔	✔	✔	✔
Kobelt	2011	[38]	Sweden	✔	✔	✔	✔	✔	✔	✔
Lindgren cluster	2009	[39, 40]	Sweden	✔	✔	✔	✔	✔	✔	✔
Maetzel	2002	[41]	Canada	✔	✔	✔	✔	–	✔	–
Marra	2007	[42]	Canada	✔	–	✔	✔	✔	✔	✔
Merkesdal	2010	[43]	Germany	✔	✔	✔	✔	✔	✔	–
Nguyen	2012	[44]	US	✔	–	✔	✔	✔	✔	–
Puolakka	2012	[45]	Finland	✔	✔	✔	✔	–	–	✔
Russell	2009	[46]	Canada	✔	✔	✔	✔	–	–	✔
Saraux	2010	[47]	France	✔	✔	✔	✔	–	–	✔
Schadlich	2005	[48]	Germany	✔	✔	✔	✔	–	✔	–
Soini	2012	[49]	Finland	✔	✔	✔	✔	✔	✔	–
Spalding	2006	[50]	US	✔	–	✔	✔	✔	✔	✔
Suka	2004	[51]	Japan	✔	–	–	–	–	–	✔
Tanno	2006	[52]	Japan	✔	✔	✔	✔	✔	✔	–
Tosh	2011	[53]	UK	✔	✔	✔	✔	–	✔	–
Tosh–Sheffield model	2011	[54]	UK	✔	✔	✔[a]	✔	✔	✔	✔
Vera-Llonch cluster	2008–2010	[55–57]	US	✔	–	✔	✔	✔	✔	✔
Wailoo	2008	[58]	US	✔	–	✔	✔	✔	✔	✔
Welsing cluster	2004–2011	[59–61]	Netherlands	✔	✔	✔	✔	–	✔	✔
Wong cluster	2002–2005	[62, 63]	US and UK	✔	✔	–	✔	✔	✔	✔

BRAM Birmingham Rheumatoid Arthritis Model, *QALY* quality-adjusted life-year, ✔ requirement met, – requirement not met

[a] Not all required information is mentioned in the referred paper. From other papers that have used/adapted the Sheffield model, we understand that all criteria can theoretically be fulfilled [15, 23, 39, 47]. However, these models were regarded as unsuitable for clustering, as each model contains different structural adaptations

variation. The adjustment of these models can be regarded as a simple adaptation, and therefore these models are not excluded.

The Birmingham Rheumatoid Arthritis Model (BRAM) was developed for the purpose of a multiple technology appraisal program, commissioned by the National Institute for Health and Care Excellence (NICE), and was designed to be flexible in changing interventions and treatment sequences in RA [18]. The individual sampling model was populated with UK data, but these can be replaced by local data. In addition, the model has an option to specify stages of early withdrawal, thereby enabling treatment discontinuation within a certain number of weeks of treatment in accordance with local clinical practice. It seems therefore

that the risk of case-mix or practice variation bias after a simple adaptation procedure is limited with this model.

Like the BRAM, the Sheffield model is a model that has been adapted several times and used for various research purposes over the years [54]. Some of the models included in this review—for example, the model by Brennan et al. [24]—are descendants of the Sheffield model but, as a result of different structural adaptations, the modelling methods have diverged, making these models not identical and therefore difficult to cluster. The general Sheffield model, as described by Tosh et al. [54], is flexible in terms of evaluating different patient populations and is able to utilize a wide range of data types. On the basis of this flexibility and the key inputs that are required by the model (initial effectiveness, HAQ progression while on treatment and length of treatment), it seems that the model is sufficiently adaptable to prevent case-mix or practice variation bias in the transferred model. However, the general model structure has been described very briefly, and therefore it is difficult to assess the complexity of the adaptation process.

The specific knock-out criterion 'technology availability' was not considered a risk for structural bias when transferring any of the models. The therapeutic technologies of interest in this case (in the treatment of RA) mainly concern drugs and medical personnel, which are not expected to differ much between the UK/Sweden and the Netherlands. The criterion 'disease spread' was not applicable to the case of RA and was therefore not considered. The result of the assessment is presented in Table 4.

3.2.3 Step 3: Which Model Is Best?

The BRAM, as well as the model by Brennan, appeared to have the highest quality according to the Philips checklist (Table 5). However, many criteria focused on the reports of the study and the model inputs, rather than the model structure. As model inputs are adaptable, the criteria looking at the model structure may be considered more relevant. Looking more specifically at the criteria of the model structure, all four remaining models were considered of good quality and suitable for simple adaptation [18, 24, 39, 54].

The choice of one of the models may eventually depend on specific model characteristics, data availability and user preferences. A comparison of modelling methods in more detail may consolidate the choice. The basic structures of the four suitable models are relatively similar; they are all individual sampling models consisting of multiple treatment lines; they use HAQ progression as the clinical outcome, they define treatment switching on the basis of the time on treatment or time to an event, and they link the quality of life to HAQ scores. However, differences are also present. For example, only the model by Lindgren et al. includes DAS in addition to HAQ scores to define its health states. The models by Lindgren et al. and Brennan et al. are based on data from observational studies, whereas the BRAM and Sheffield models are based on data derived from clinical trials. Brennan et al. include only direct costs related to treatment, whereas the other models also include costs of resource use linked to disease severity. Both Lindgren et al. and the BRAM model apply mortality rates on the basis of disease severity, while the other models apply standard life tables. In the Netherlands, DAS, next to HAQ scores, is considered to be a relevant outcome measure, and patient-level observational data on these outcomes and on quality of life (QoL) outcomes are available. Therefore, the model by Lindgren et al. seems to be a good choice for the Netherlands.

Table 4 Results of applying the specific knock-out criteria

Model	Specific knock-out criteria		
	Practice variation	Technology availability	Case mix
BRAM cluster [17–22]	Good fit	Good fit	Good fit
Brennan [24]	Good fit	Good fit	Good fit
Finckh [29]	Imperfect fit	Good fit	Good fit
Kobelt [38]	Imperfect fit	Good fit	Good fit
Lindgren cluster [39, 40]	Good fit	Good fit	Good fit
Tosh–Sheffield model [54]	Good fit	Good fit	Good fit

The criterion 'disease spread' was left out of the table, as this is not applicable for rheumatoid arthritis models

BRAM Birmingham Rheumatoid Arthritis Model

4 Discussion

In this study, we performed a stepwise process for the selection of a health economic evaluation model that is suitable for simple adaptation. We found that six models fulfilled the pre-specified minimal methodological and structural requirements, of which four models seemed to have the best fit with the Dutch health care setting, and all four were of good quality. These models were therefore regarded as suitable to transfer by a simple adaptation.

To our knowledge, no examples of similar model selection exercises are available in the literature. Although it is common practice that health economic evaluation models are newly developed for specific decision problems, there are also many examples of models that have been adapted and re-used for multiple purposes within multiple countries. Often this occurred within research groups who developed a model and used this model multiple times to answer different research questions (such as the studies performed by Kobelt et al. [32–37]). Apart from

Table 5 Results of applying the Philips quality checklist

Philips criteria		BRAM[a] [17–22]	Brennan [24]	Lindgren [39, 40]	Sheffield [54]
Structure					
S1	Is there a clear statement of the decision problem?	Yes	Yes	Yes	NA
	Is the objective specified and consistent with the stated decision problem?	Yes	Yes	Yes	NA
	Is the primary decision maker specified?	Yes	Yes	No	NA
S2	Is the perspective of the model stated clearly?	Yes	Yes	Yes	NA
	Are the model inputs consistent with the stated perspective?	Yes	Yes	Yes	NA
	Has the scope of the model been stated and justified?	Yes	Yes	Yes	NA
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Yes	Yes	Yes	NA
S3	Has the evidence regarding the model structure been described?	Yes	Yes	Yes	Yes
	Are the sources of data used to develop the structure of the model specified?	No	No	No	No
	Are the causal relationships described by the model structure justified appropriately?	Yes	Yes	Yes	Yes
S4	Are the structural assumptions transparent and justified?	Yes	Yes	Yes	Yes/No
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Yes	Yes	Yes	Yes
S5	Is there a clear definition of the options under evaluation?	Yes	Yes	Yes	NA
	Have all feasible and practical options been evaluated?	Yes	Yes	Yes	NA
	Is there justification for the exclusion of feasible options?	NA	NA	NA	NA
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	Yes	Yes	Yes	Yes
S7	Is the time horizon of the model sufficient to reflect all important differences between options?	Yes	Yes	Yes	Yes
	Is the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	Yes	Yes	Yes	Yes
S8	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	Yes	Yes	Yes	Yes
S9	Is the cycle length defined and justified in terms of the natural history of disease?	NA	NA	NA	NA
Data					
D1	Are the data identification methods transparent and appropriate given the objectives of the model?	Yes	Yes	Yes	NA
	Where choices have been made between data sources, are these justified appropriately?	Yes	Yes	Yes	NA
	Has particular attention been paid to identifying data for the important parameters in the model?	Yes	Yes	Yes	NA
	Has the process of selecting key parameters been justified and systematic methods used to identify the most appropriate data?	Yes	Yes	Yes	NA
	Has the quality of the data been assessed appropriately?	Yes	Yes	Yes	NA
	Where expert opinion has been used, are the methods described and justified?	NA	NA	NA	NA
D2	Is the pre-model data analysis methodology based on justifiable statistical and epidemiological techniques?	Yes	Yes	Yes/No	NA
D2a	Is the choice of baseline data described and justified?	Yes	Yes	Yes	NA
	Are transition probabilities calculated appropriately?	Yes	Yes	Yes	NA
	Has a half cycle correction been applied to both cost and outcome?	NA	NA	NA	NA
	If not, has this omission been justified?	NA	NA	NA	NA

Table 5 continued

Philips criteria		BRAM[a] [17–22]	Brennan [24]	Lindgren [39, 40]	Sheffield [54]
D2b	If relative treatment effects have been derived from trial data, have they been synthesized using appropriate techniques?	Yes/No	Yes	Yes/No	NA
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	Yes	Yes	Yes	Yes
	Have alternative extrapolation assumptions been explored through sensitivity analysis?	Yes	Yes	Yes/No	NA
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	Yes	Yes	Yes	Yes
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	No	Yes	Yes	NA
D2c	Are the utilities incorporated into the model appropriate?	Yes	Yes	Yes	NA
	Is the source for the utility weights referenced?	Yes	Yes	Yes	NA
	Are the methods of derivation for the utility weights justified?	Yes	Yes	Yes	NA
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	Yes	Yes/No	No	NA
	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate?)	NA	NA	Yes/No	NA
	Is the process of data incorporation transparent?	Yes	Yes	Yes	NA
	If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	Yes/No	No	No	NA
	If data have been incorporated as distributions, is it clear that second-order uncertainty is reflected?	Yes	Yes	Yes	NA
D4	Have the four principal types of uncertainty been addressed?	Yes	Yes	Yes	NA
	If not, has the omission of particular forms of uncertainty been justified?	NA	NA	NA	NA
D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	Yes	Yes	Yes	NA
D4b	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	Yes	Yes	Yes	NA
D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	No	Yes	Yes	NA
D4d	Are the methods of assessment of parameter uncertainty appropriate?	Yes	Yes	Yes	NA
	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	Yes/No	NA	No	NA
Consistency					
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	No	No	No	No
C2	Are the conclusions valid given the data presented?	Yes	Yes	Yes	NA
	Are any counterintuitive results from the model explained and justified?	NA	NA	NA	NA
	If the model has been calibrated against independent data, have any differences been explained and justified?	NA	NA	NA	NA
	Have the results of the model been compared with those of previous models and any differences in results explained?	Yes	Yes	No	NA

BRAM Birmingham Rheumatoid Arthritis Model, *NA* criterion not applicable, *Yes/No* answer is debatable

[a] Malottki et al., 2011 [17], Barton et al., 2011 [19]

the model developers, parties such as industry also used existing models to perform health economic research (for example, the Center for Outcomes Research & Evaluation [CORE] diabetes model) [64]. Especially in therapeutic areas such as RA and diabetes, where high-quality health economic evaluation models have been developed over the years, it can be a challenge to develop a model that approaches the quality of some of the existing models. Adapting a good-quality model may be an efficient use of resources, provided that the model fits well for the jurisdiction of interest. We believe that a systematic and explicit approach for model selection can contribute to the validity of the study results when existing models are adapted to study new decision problems.

The methodology we chose for this exercise was predefined and was based mainly on the ISPOR Good Research Practices on the Transferability of Economic Evaluations Across Jurisdictions [2] and on the decision chart of Welte [4]. The decision chart of Welte starts with three 'general knock-out criteria', which concern the evaluated technology, the comparator and the study quality. The criterion of study quality was adopted in our approach but was, for reasons of logic and efficiency, redefined to be our last step. Because Welte did not specify the method for the assessment of study quality, we proposed to apply the Philips checklist, which is quite elaborate. The general knock-out criteria 'evaluated technology' and 'comparator' were not adopted in our approach for the following reasons. First, models that include a treatment or comparator different from that in the research question of interest may be perfectly suitable to be (simply) adapted to include the treatment or comparator of interest. Second, if only models that include a specific treatment or comparator are to be considered, such models can be selected by specifying the inclusion and exclusion criteria in the systematic literature search. Third, if the treatment or comparator of interest requires specific methodological/structural model features, these can be specified in the 'minimal methodological requirements' of step 1 of the selection process.

Conceptual validity is of major importance when selecting a health economic evaluation model, as this secures a fair representation of the treatment and contextual reality. For the case of RA, the OMERACT reference case was regarded as the most suitable guidance for defining the requirements to determine conceptual validity and therefore served as the basis for these criteria. However, the criteria may have been too strict because in some cases, the excluded models met most of the criteria and the failed criterion might have been overcome if the model had been transferred (e.g. it probably would not have been very complicated to adapt the model by Welsing et al. [60] to include mortality). We also encountered one study that reported disability-adjusted life-years (DALYs) instead of QALYs, and it was (for this and other reasons) excluded [51]. Although DALYs are not mentioned by OMERACT as a relevant outcome, DALYs can be a valid outcome measure for health economic evaluations.

For future application, it is advisable to specify the minimal methodological and structural requirements based on the available evidence of good methodological practice in the therapeutic area of interest. The OMERACT reference case serves as a good example of such evidence, and it would be useful if such exercises were also undertaken in other disease areas to derive consensus-based criteria, as has previously been proposed within the field of osteoarthritis [65]. If such a reference case is not available, it may be worthwhile to search for other sources of guidance—for example, from methodological publications on health economic modelling or from previous literature reviews of models in the therapeutic area of interest. Also, it may be necessary to account for the specific research question when specifying the requirements—for example, if a certain intervention requires specific structural or functional modelling options. In addition, the minimal methodological and structural requirements should be customized to the jurisdiction to which the model will be transferred, as many jurisdictions have their own guidelines when it comes to health economic research (e.g. on the time horizon of the analysis), and the model should be able to produce all required results. Naturally, the outcome of the selection process also very much depends on the jurisdiction of interest—for example, where Swedish models may show a relatively good fit when transferred to the Netherlands, these models will probably show a poor fit when transferred to a jurisdiction where health care practice is very different. Had we taken another country as an illustrative case, other models would possibly have passed our selection criteria, resulting in a different set of selected models.

It is questionable whether applying a different order of selection steps is more appropriate. We could have reversed steps 2 and 3, thereby applying the Philips checklist prior to the specific knock-out criteria for model fit. However, the Philips checklist does not have a clear cut-off point between good and poor quality, making model selection difficult. In addition, on the basis of our experience, it takes much more time to go through the Philips checklist (±3 h) than to go through the specific knock-out criteria (±30 min), and it is therefore more pragmatic to apply the Philips checklist as a last step. Another concern is that the total time required to conduct this selection procedure might be longer than that needed to develop a new model. However, even if the modelling time is shorter (which is almost impossible to affirm prior to conducting the selection procedure), knowledge of the existing models and previously applied methodology is valuable and can be considered a starting point for developing a new (preferably superior) model. In any case, the currently proposed procedure is helpful, especially in justifying the researcher decision on whether or not a new model should be developed. We thus conclude that our proposed selection process is appropriate and feasible.

A limitation of the current exercise is that it was not possible to assess the computerized model validity, as the modelling codes were not published. This reflects a significant problem for researchers who want to use an existing health economic evaluation model—namely, that most models are not easily available to others than those who developed it. Generally, models are regarded as the intellectual property of the developer, which is

understandable given that it may take months to years of work for the models to be developed. This makes model selection complicated, as detailed descriptions of the model are not always provided in the papers and the model software is usually not published. Without access to the models, the assessment of the model fit is also subject to uncertainty, as it is impossible to describe all methodological details in a paper. Furthermore, once the selection process is completed, the possibility to transfer the model depends on the willingness of the model owner to provide access to the model and to permit the use of it.

Availability of (or access to) the model is required in order to successfully transfer health economic evaluation models. Broad availability of models could improve the efficiency of health economic research and the quality of the models and the resulting health economic evidence, as researchers could cooperate and provide feedback on each other's work. In addition, it would become much easier to assess the validity of health economic evidence, and it would make health economic studies more reproducible—a general prerequisite in scientific research [66]. If this ideal situation is established, transferring models will be highly possible in the context of scientific cooperation, under licensing conditions or by other financial arrangements. Some models are already open source, such as the Evidence-Based Medicine Integrator simulator for type 2 diabetes mellitus [67] and the Comprehensive Discrete Event Simulation Model for Major Depressive Disorder, developed by Toumi et al. [68].

5 Conclusion

Overall, we conclude that, prior to development of a new health economic evaluation model for a specific decision problem, application of a systematic approach to identification of existing models—and selection of a qualitative model that is transferable—is feasible and is expected to contribute to the efficiency, quality and validity of the resulting health economic evidence.

Disclosures The author contributions to this manuscript were as follows:

HvH: study rationale and design, execution of systematic literature search, study selection, execution of stepwise selection, interpretation and reflection, writing of the manuscript.

JS: study rationale and design, study selection, verification of stepwise selection, interpretation and reflection, review of the manuscript, overall guarantor of the study.

ATD: study selection, feedback on study execution and interpretation, review of the manuscript.

AB: feedback on study execution and interpretation, review of the manuscript.

HvH is employed by AstraZeneca and affiliated with the Erasmus University Rotterdam Institute of Health Policy & Management by means of a PhD hospitality agreement. JS's supervision of HvH is compensated by AstraZeneca on the basis of an unrestricted supervision grant to his institution. In the field of rheumatoid arthritis, JS has received research grants from Pfizer. AB has received research grants from Merck, AbbVie and Amgen. HvH, JS, ATD and AB declare no conflict of interest regarding the topics studied and the results presented in this publication.

References

1. Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, et al. Generalisability in economic evaluation studies in healthcare: a review and case studies. Health Technol Assess. 2004;8(49):iii,iv, 1–192.
2. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. Value Health. 2009;12(4):409–18.
3. Goeree R, He J, O'Reilly D, Tarride JE, Xie F, Lim M, et al. Transferability of health technology assessments and economic evaluations: a systematic review of approaches for assessment and application. Clinicoecon Outcomes Res. 2011;3:89–104.
4. Welte R, Feenstra T, Jager H, Leidl R. A decision chart for assessing and improving the transferability of economic evaluation results between countries. Pharmacoeconomics. 2004;22(13):857–76.
5. Alamanos Y, Drosos AA. Epidemiology of adult rheumatoid arthritis. Autoimmun Rev. 2005;4(3):130–6.
6. Eriksson JK, Neovius M, Ernestam S, Lindblad S, Simard JF, Askling J. Incidence of rheumatoid arthritis in Sweden—a nationwide population-based assessment of incidence, its determinants, and treatment penetration. Arthritis Care Res (Hoboken). 2013;65:870–8.
7. Strand V, Singh JA. Improved health-related quality of life with effective disease-modifying antirheumatic drugs: evidence from randomized controlled trials. Am J Manag Care. 2008;14(4):234–54.
8. Fautrel B. Economic benefits of optimizing anchor therapy for rheumatoid arthritis. Rheumatology (Oxford). 2012 Jun;51 Suppl 4:iv21–6.
9. Maetzel A, Tugwell P, Boers M, Guillemin F, Coyle D, Drummond M, et al. Economic evaluation of programs or interventions in the management of rheumatoid arthritis: defining a consensus-based reference case. J Rheumatol. 2003;30(4):891–6.
10. Bansback N, Ara R, Karnon J, Anis A. Economic evaluations in rheumatoid arthritis: a critical review of measures used to define health states. Pharmacoeconomics. 2008;26(5):395–408.
11. Bansback NJ, Regier DA, Ara R, Brennan A, Shojania K, Esdaile JM, et al. An overview of economic evaluations for drugs used in rheumatoid arthritis: focus on tumour necrosis factor-alpha antagonists. Drugs. 2005;65(4):473–96.
12. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force—6. Value Health. 2012;15(6):835–42.
13. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24(4):355–71.
14. 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. 2011;29(4):633–41.

15. Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. Ann Rheum Dis. 2005;64(7):995–1002.
16. 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 Inflamm. 2011;2011:727634.
17. Malottki K, Barton P, Tsourapas A, Uthman AO, Liu Z, Routh K, et al. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation. Health Technol Assess. 2011;15(14):1–278.
18. Barton P. Development of the Birmingham Rheumatoid Arthritis Model: past, present and future plans. Rheumatology (Oxford). 2011;50 Suppl 4:iv32–8.
19. 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(11):iii, 1–91.
20. Chen YF, Jobanputra P, Barton P, Jowett S, Bryan S, Clark W, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. Health Technol Assess. 2006;10(42):iii, iv, xi–xiii, 1–229.
21. Clark W, Jobanputra P, Barton P, Burls A. The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis. Health Technol Assess. 2004;8(18):iii, iv, ix–x, 1–105.
22. Jobanputra P, Barton P, Bryan S, Burls A. The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation. Health Technol Assess. 2002;6(21):1–110.
23. Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. Rheumatology (Oxford). 2004;43(1):62–72.
24. 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.
25. Chiou CF, Choi J, Reyes CM. Cost-effectiveness analysis of biological treatments for rheumatoid arthritis. Expert Rev Pharmacoecon Outcomes Res. 2004;4(3):307–15.
26. Choi HK, Seeger JD, Kuntz KM. A cost effectiveness analysis of treatment options for methotrexate-naive rheumatoid arthritis. J Rheumatol. 2002;29(6):1156–65.
27. 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.
28. Diamantopoulos A, Benucci M, Capri S, Berger W, Wintfeld N, Giuliani G, et al. Economic evaluation of tocilizumab combination in the treatment of moderate-to-severe rheumatoid arthritis in Italy. J Med Econ. 2012;15(3):576–85.
29. Finckh A, Bansback N, Marra CA, Anis AH, Michaud K, Lubin S, et al. Treatment of very early rheumatoid arthritis with symptomatic therapy, disease-modifying antirheumatic drugs, or biologic agents: a cost-effectiveness analysis. Ann Intern Med. 2009;151(9):612–21.
30. Hallinen TA, Soini EJ, 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 (Oxford). 2010;49(4):767–77.
31. Kielhorn A, Porter D, Diamantopoulos A, Lewis G. UK cost–utility analysis of rituximab in patients with rheumatoid arthritis that failed to respond adequately to a biologic disease-modifying antirheumatic drug. Curr Med Res Opin. 2008;24(9):2639–50.
32. Lekander I, Borgstrom F, Svarvar P, Ljung T, Carli C, van Vollenhoven RF. Cost-effectiveness of real-world infliximab use in patients with rheumatoid arthritis in Sweden. Int J Technol Assess Health Care. 2010;26(1):54–61.
33. Kobelt G, Jonsson L, Lindgren P, Young A, Eberhardt K. Modeling the progression of rheumatoid arthritis: a two-country model to estimate costs and consequences of rheumatoid arthritis. Arthritis Rheum. 2002;46(9):2310–9.
34. Kobelt G, Lindgren P, Young A. Modelling the costs and effects of leflunomide in rheumatoid arthritis. Eur J Health Econ. 2002;3(3):180–7.
35. Kobelt G, Jonsson L, Young A, Eberhardt K. The cost-effectiveness of infliximab (remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study. Rheumatology (Oxford). 2003;42(2):326–35.
36. Kobelt G, Lindgren P, Lindroth Y, Jacobson L, Eberhardt K. Modelling the effect of function and disease activity on costs and quality of life in rheumatoid arthritis. Rheumatology (Oxford). 2005;44(9):1169–75.
37. Kobelt G, Lindgren P, Singh A, Klareskog L. Cost effectiveness of etanercept (Enbrel) in combination with methotrexate in the treatment of active rheumatoid arthritis based on the TEMPO trial. Ann Rheum Dis. 2005;64(8):1174–9.
38. Kobelt G, Lekander I, Lang A, Raffeiner B, Botsios C, Geborek P. Cost-effectiveness of etanercept treatment in early active rheumatoid arthritis followed by dose adjustment. Int J Technol Assess Health Care. 2011;27(3):193–200.
39. Lindgren P, Geborek P, Kobelt G. Modeling the cost-effectiveness of treatment of rheumatoid arthritis with rituximab using registry data from southern Sweden. Int J Technol Assess Health Care. 2009;25(2):181–9.
40. Kobelt G, Lindgren P, Geborek P. Costs and outcomes for patients with rheumatoid arthritis treated with biological drugs in Sweden: a model based on registry data. Scand J Rheumatol. 2009;38(6):409–18.
41. Maetzel A, Strand V, Tugwell P, Wells G, Bombardier C. Cost effectiveness of adding leflunomide to a 5-year strategy of conventional disease-modifying antirheumatic drugs in patients with rheumatoid arthritis. Arthritis Rheum. 2002;47(6):655–61.
42. Marra CA, Marion SA, Guh DP, Najafzadeh M, Wolfe F, Esdaile JM, et al. Not all "quality-adjusted life years" are equal. J Clin Epidemiol. 2007;60(6):616–24.
43. Merkesdal S, Kirchhoff T, Wolka D, Ladinek G, Kielhorn A, Rubbert-Roth A. Cost-effectiveness analysis of rituximab treatment in patients in Germany with rheumatoid arthritis after etanercept-failure. Eur J Health Econ. 2010;11(1):95–104.
44. Nguyen CM, Bounthavong M, Mendes MA, Christopher ML, Tran JN, Kazerooni R, et al. Cost utility of tumour necrosis factor-alpha inhibitors for rheumatoid arthritis: an application of Bayesian methods for evidence synthesis in a Markov model. Pharmacoeconomics. 2012;30(7):575–93.
45. 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.
46. 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.
47. 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.

48. Schadlich PK, Zeidler H, Zink A, Gromnica-Ihle E, Schneider M, Straub C, et al. Modelling cost effectiveness and cost utility of sequential DMARD therapy including leflunomide for rheumatoid arthritis in Germany: II. The contribution of leflunomide to efficiency. Pharmacoeconomics. 2005;23(4):395–420.
49. Soini EJ, Hallinen TA, Puolakka K, Vihervaara V, Kauppi MJ. Cost-effectiveness of adalimumab, etanercept, and tocilizumab as first-line treatments for moderate-to-severe rheumatoid arthritis. J Med Econ. 2012;15(2):340–51.
50. Spalding JR, Hay J. Cost effectiveness of tumour necrosis factor-alpha inhibitors as first-line agents in rheumatoid arthritis. Pharmacoeconomics. 2006;24(12):1221–32.
51. Suka M, Yoshida K. Cost effectiveness of leflunomide in the treatment of rheumatoid arthritis in Japan. Expert Rev Pharmacoecon Outcomes Res. 2004;4(6):617–22.
52. Tanno M, Nakamura I, Ito K, Tanaka H, Ohta H, Kobayashi M, et al. Modeling and cost-effectiveness analysis of etanercept in adults with rheumatoid arthritis in Japan: a preliminary analysis. Mod Rheumatol. 2006;16(2):77–84.
53. Tosh JC, Wailoo AJ, Scott DL, Deighton CM. Cost-effectiveness of combination nonbiologic disease-modifying antirheumatic drug strategies in patients with early rheumatoid arthritis. J Rheumatol. 2011;38(8):1593–600.
54. Tosh J, Brennan A, Wailoo A, Bansback N. The Sheffield Rheumatoid Arthritis Health Economic Model. Rheumatology (Oxford). 2011;50 Suppl 4:iv26–31.
55. Yuan Y, Trivedi D, Maclean R, Rosenblatt L. Indirect cost-effectiveness analyses of abatacept and rituximab in patients with moderate-to-severe rheumatoid arthritis in the United States. J Med Econ. 2010;13(1):33–41.
56. Vera-Llonch M, Massarotti E, Wolfe F, Shadick N, Westhovens R, Sofrygin O, et al. Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to tumor necrosis factor-alpha antagonists. J Rheumatol. 2008;35(9):1745–53.
57. Vera-Llonch M, Massarotti E, Wolfe F, Shadick N, Westhovens R, Sofrygin O, et al. Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to methotrexate. Rheumatology (Oxford). 2008;47(4):535–41.
58. Wailoo AJ, Bansback N, Brennan A, Michaud K, Nixon RM, Wolfe F. Biologic drugs for rheumatoid arthritis in the Medicare program: a cost-effectiveness analysis. Arthritis Rheum. 2008;58(4):939–46.
59. Welsing PM, Severens JL, Hartman M, van Gestel AM, van Riel PL, Laan RF. The initial validation of a Markov model for the economic evaluation of (new) treatments for rheumatoid arthritis. Pharmacoeconomics. 2006;24(10):1011–20.
60. Welsing PM, Severens JL, Hartman M, van Riel PL, Laan RF. Modeling the 5-year cost effectiveness of treatment strategies including tumor necrosis factor-blocking agents and leflunomide for treating rheumatoid arthritis in the Netherlands. Arthritis Rheum. 2004;51(6):964–73.
61. Schipper LG, Kievit W, den Broeder AA, van der Laar MA, Adang EM, Fransen J, et al. Treatment strategies aiming at remission in early rheumatoid arthritis patients: starting with methotrexate monotherapy is cost-effective. Rheumatology (Oxford). 2011;50(7):1320–30.
62. Barbieri M, Wong JB, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. Pharmacoeconomics. 2005;23(6):607–18.
63. Wong JB, Singh G, Kavanaugh A. Estimating the cost-effectiveness of 54 weeks of infliximab for rheumatoid arthritis. Am J Med. 2002;113(5):400–8.
64. Beaudet A, Palmer JL, Timlin L, Wilson B, Bruhn D, Boye KS, et al. Cost–utility of exenatide once weekly compared with insulin glargine in patients with type 2 diabetes in the UK. J Med Econ. 2011;14(3):357–66.
65. Hiligsmann M, Cooper C, Arden N, Boers M, Branco JC, Luisa Brandi M, et al. Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). Semin Arthritis Rheum. 2013;43(3):303–13.
66. The Royal Society. Science as an open enterprise: open data for open science; 2012. http://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf.
67. Blum S, Vardi M, Brown JB, Russell A, Milman U, Shapira C, et al. Vitamin E reduces cardiovascular disease in individuals with diabetes mellitus and the haptoglobin 2-2 genotype. Pharmacogenomics. 2010;11(5):675–84.
68. Toumi M, Antonanzas F, Hakkaart L, Lam RW, McCrone P, Persson U, et al. Comprehensive discrete event simulation model for the evaluation of health care technologies in depression. Value Health. 2012;15(7):A282.