



Key data elements for use in cost-utility modeling of biological treatments for rheumatoid arthritis

Michael L. Ganz, Brian Bekker Hansen, Xavier Valencia & Martin Strandberg-Larsen

To cite this article: Michael L. Ganz, Brian Bekker Hansen, Xavier Valencia & Martin Strandberg-Larsen (2015) Key data elements for use in cost-utility modeling of biological treatments for rheumatoid arthritis, Journal of Medical Economics, 18:5, 366-375, DOI: [10.3111/13696998.2014.1001848](https://doi.org/10.3111/13696998.2014.1001848)

To link to this article: <https://doi.org/10.3111/13696998.2014.1001848>



Published online: 19 Jan 2015.



Submit your article to this journal [↗](#)



Article views: 469



View related articles [↗](#)



View Crossmark data [↗](#)

Original article

Key data elements for use in cost-utility modeling of biological treatments for rheumatoid arthritis

Michael L. Ganz

Evidera, Lexington, MA, USA

Brian Bekker Hansen

Novo Nordisk A/S, Bagsværd, Denmark

Xavier Valencia

Bristol Myers Squibb, Princeton, NJ, USA

Martin Strandberg-Larsen

Novo Nordisk A/S, Bagsværd, Denmark

Address for correspondence:

Michael L. Ganz, MS, PhD, Evidera 430 Bedford Street, Suite 300, Lexington, MA 02420, USA.
Tel: +1-781-960-0314;
michael.ganz@evidera.com

Keywords:

Rheumatoid arthritis – Cost-utility – Economic models
– Clinical trial design – Pharmacoeconomics

Accepted: 19 December 2014; published online: 19 January 2015
Citation: J Med Econ 2015; 18:366–75

Abstract**Objectives:**

Economic evaluation is becoming more common and important as new biologic therapies for rheumatoid arthritis (RA) are developed. While much has been published about how to design cost-utility models for RA to conduct these evaluations, less has been written about the sources of data populating those models. The goal is to review the literature and to provide recommendations for future data collection efforts.

Methods:

This study reviewed RA cost-utility models published between January 2006 and February 2014 focusing on five key sources of data (health-related quality-of-life and utility, clinical outcomes, disease progression, course of treatment, and healthcare resource use and costs). It provided recommendations for collecting the appropriate data during clinical and other studies to support modeling of biologic treatments for RA.

Results:

Twenty-four publications met the selection criteria. Almost all used two steps to convert clinical outcomes data to utilities rather than more direct methods; most did not use clinical outcomes measures that captured absolute levels of disease activity and physical functioning; one-third of them, in contrast with clinical reality, assumed zero disease progression for biologic-treated patients; little more than half evaluated courses of treatment reflecting guideline-based or actual clinical care; and healthcare resource use and cost data were often incomplete.

Conclusions:

Based on these findings, it is recommended that future studies collect clinical outcomes and health-related quality-of-life data using appropriate instruments that can convert directly to utilities; collect data on actual disease progression; be designed to capture real-world courses of treatment; and collect detailed data on a wide range of healthcare resources and costs.

Introduction

Rheumatoid arthritis (RA) is a chronic disease characterized by progressive and irreversible joint damage that negatively impacts patients' physical health, functional status, independence, and emotional well-being. Current treatment options include synthetic, or traditional, disease modifying anti-rheumatic drugs (sDMARDs), which are used in essentially all patients with RA, biologic disease modifying anti-rheumatic drugs (bDMARDs), and adjunctive therapies such as analgesics, non-steroidal anti-inflammatory drugs, and corticosteroids to address symptoms. The availability of bDMARDs has provided additional treatment options to patients who, because they no longer respond to sDMARDs,

have experienced adverse events or toxicity, or have exhausted their treatment options^{1,2}. These new and effective biological therapies clearly benefit patients, but, just as their clinical value must be assessed, their economic value should be assessed as well, particularly because of their high costs. However, because clinical trials, and prospective observational studies, may not collect information for long enough periods of time to assess long-term outcomes and because they may not capture all necessary clinical and economic information, clinical-economic models are used to synthesize data from disparate sources to simulate and extrapolate long-term costs and clinical outcomes³. For such models to be useful decision aids, they need to validly capture the costs and outcomes of different RA therapies. The absence of a common or core set of measurements collected during clinical trials, or obtained from other study designs, that may serve as data sources for the models can complicate the task of comparing models and their results and contributes to the variability in the types of models and data used.

In 2003, the Outcome Measures in Rheumatology (OMERACT) Economics Working Group published a proposal for the reference case of an economic evaluation in RA^{4,5}. The OMERACT proposal included 12 recommendations that were grouped into four broad categories covering outcome measures, comparators, modeling techniques, and costs. Their recommendations focused on model design choices, but not on the sources of the data used in the model (i.e., what comes before the model design stage). In 2005, Drummond *et al.*⁶ compared economic models of etanercept and infliximab^{7–11} that were published between 2002–2005. They evaluated the models according to 23 criteria that were grouped into two broad categories that were responsible for differences in the outcomes/results of those models, structural features and input parameters, with input data sometimes being the major driver of those different results. Their recommendations for future work did not focus on data sources or data collection design issues. The National Institute for Health and Care Excellence (NICE) in the UK published a systematic review of the clinical effectiveness and performed economic evaluations of adalimumab, etanercept, and infliximab in 2006 and expanded their previous review and evaluations by including rituximab and abatacept in 2011^{12,13}. These NICE assessments also contained an independent economic evaluation of the bDMARDs under consideration, which was based on the Birmingham Rheumatoid Arthritis Model (BRAM).

The OMERACT, Drummond *et al.*⁶, and NICE publications discussed how model design and selection of input data are closely related to model validity and how model results can be influenced by those characteristics. Much has already been written about RA model design, the importance of these attributes, and how they determine model validity; therefore, this is not the goal of this article.

Much less, however, has been written about how the data themselves and their sources can influence a model's validity. The purpose of this article is to provide recommendations to enhance the design of clinical and other studies so that a wider range of outcomes and economic data can be collected to support future economic analyses. This is particularly important in light of the high costs of biologic therapies. The discussion and recommendations are based on a literature review of published RA economic models.

Methods

We reviewed economic models of RA published in the peer-reviewed English-language literature between January 1, 2006 and February 18, 2014. We searched the Medline and Embase databases as well as the Cochrane Library using pre-defined search algorithms (see the Appendix). Specifically, we identified studies that presented information on decision-analytic models (decision trees, Markov models, discrete event simulation models, or other simulation models) comparing at least two treatment strategies for patients with RA, one of which must have included a biologic treatment, that reported (1) costs and outcomes associated with each treatment, (2) cost-effectiveness results in terms of costs per QALY (i.e., cost-utility results), and (3) the assumptions and technical details of the model. Economic evaluations that assess costs per clinical outcome such as cost per ACR20 response achieved or cost per patient in remission are referred to as cost-effectiveness analyses (CEA), while those that assess cost per QALYs are referred to as cost-utility analyses (CUA). Although robust CEAs have been performed, and may be preferred to CUAs due to the limitations of using utilities and QALYs to measure clinical benefits¹⁴, the focus here, however, is on CUAs because of the reliance on CUAs for many HTA processes and local requirements.

We did not, *a priori*, exclude studies that reported on the costs and outcomes directly observed from a trial or observational database. We reviewed the full text of all sources whose titles and/or abstracts suggested that they would fulfill our inclusion and exclusion criteria. Following full text review, we extracted information about each model's design (model type, health states, time horizon, and cycle length), details on the treatment arms/strategies, data sources including how clinical outcome measures were converted to utilities, characteristics of the hypothetical patients, how disease progression was modeled, and which healthcare resources and costs were included in each model's base case. The literature search and the data extraction activities were conducted by members of our research team. All abstracts and articles that were initially marked for exclusion were verified by the lead author (MLG), as were all of the extracted data.

We developed a list of five broad key categories of model inputs that are linked to model design and that can have substantial influence on model results through the numerator and denominator of the incremental cost-utility ratio: HRQoL measures, clinical outcomes measures, disease progression, course of treatment, and healthcare resource use and costs⁴⁻⁶. In the sections that follow, we discuss how these categories of inputs were used in the models we reviewed along with gaps we have identified and we offer recommendations for appropriate data capture in a variety of settings that can be used to support clinical and economic modeling of biologic treatments for RA.

Literature review results

We identified 1421 studies and retrieved 71 for full-text review following abstract screening, 24 of which met our inclusion criteria (see Figure 1 for further study attrition information)¹⁵⁻³⁶. Half of the studies ($n = 14$) described discrete event simulation (DES) or individual patient simulation models, while the remaining described Markov cohort ($n = 9$) or decision tree ($n = 1$) models. The perspectives of 11 countries are represented with one model each from Canada, China, Colombia, Germany, Italy, and Japan, two models each from Finland, the Netherlands, and the UK, four models from Sweden, and eight models from the US. Most ($n = 13$) of

the models (and all of the ones from the US) used the perspective of a third-party payer (i.e., economic results only included direct medical costs). Models from the UK (as well as those from Canada, Germany, and Italy) used the perspective of each country's national health service.

Six studies evaluated biologic therapies in patients with early RA^{16,18,21,28,30,32}, and 16 studies evaluated biologic therapies in patients with late (or likely late) RA^{15,17,19,20,24-27,29,31-37}. We categorized stage of RA (early or late) according to the terms used by the authors (models focusing on moderate-to-severe patients who did not respond to previous treatments were classified as 'likely late' RA). Nine models, two for early RA^{16,28} and seven for late RA^{20,24,25,27,29,35,38}, assessed different treatment strategies (sequences) comprised of the same biologic therapies^{25,27,36}. Table 1 provides information on the models included in this review.

Clinical outcomes measures

Models require measures that can reliably and validly measure response to therapy and classify patients into health states for computing costs and QALYs. The three most commonly used clinical outcomes measures are the American College of Rheumatology (ACR) 20/50/70 response criteria, the DAS28, and the HAQ-DI. The ACR20/50/70 criteria measure response to treatment according to 20%, 50%, or 70% improvements in the number of tender and swollen joints plus improvements

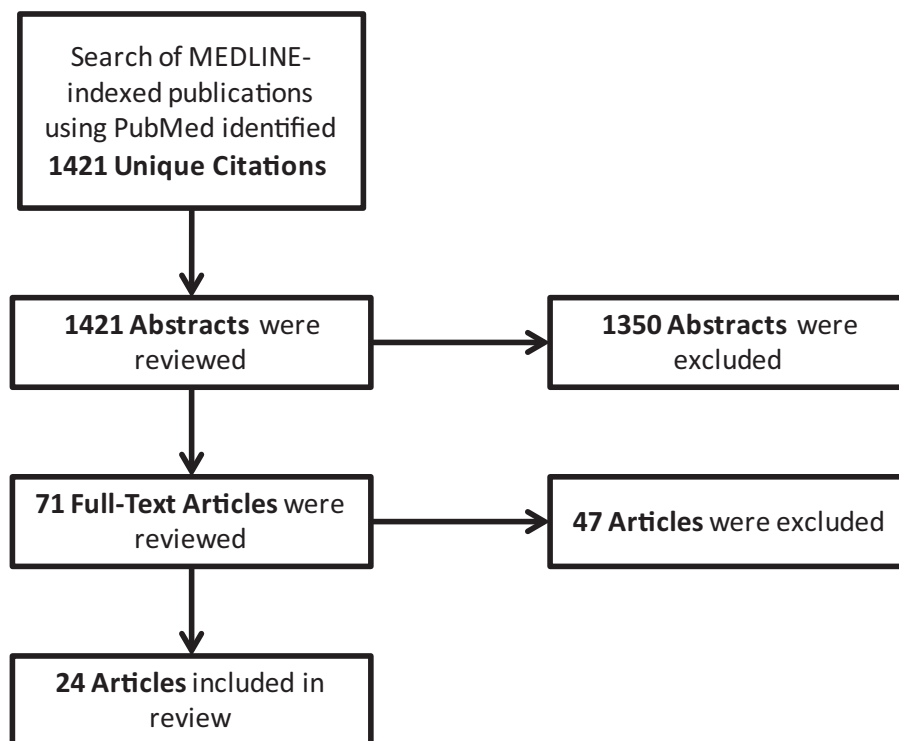


Figure 1. Study attrition of the literature review.

Table 1. Summary of reviewed models.

Citation (Country)	Cycle length, time horizon	Model type, disease description	Biologic course of treatment, treatments evaluated*	Health state utilities calculation	Clinical effectiveness measures	Disease (HAQ-DI) progression**	Healthcare resource use and costs
Brennan <i>et al.</i> ¹⁵ (UK)	6 months, Lifetime	Microsimulation, Late RA	ADA, ETA, IFX as a class	HAQ-DI, SF-6D → EQ-5D	EULAR/DAS28	+ sDMARDs	Direct medical
Davies <i>et al.</i> ¹⁶ (US)	6 months, Lifetime	Microsimulation, Early RA	ADA; ETA; IFX; ADA → ETA	ACR → HAQ-DI → HUI-3	ACR20/50/70	+ responders	Direct medical, Indirect
Diamantopoulos <i>et al.</i> ¹⁷ (Italy)	6 months, Lifetime	Microsimulation, Moderate, severe RA after sDMARDs	TOC; ETA; IFX	ACR → HAQ-DI → EQ-5D	ACR20/50/70	++ non-responders 0 biologics – TOC + palliative care	Direct medical and non-medical
Finckh <i>et al.</i> ¹⁸ (US)	6 months, Lifetime	Microsimulation, Early RA	Early biologic strategy	Radiographic damage → HAQ-DI → EQ-5D	Radiographic damage, HAQ-DI	+	Direct medical, Indirect
Hallinen <i>et al.</i> ¹⁹ (Finland)	6 months, Lifetime	Microsimulation, Late RA	ABA; ADA; ETA; IFX; RTX	ACR → HAQ-DI → HUI-3	ACR20/50/70	NR	Direct medical and non-medical
Kielhorn <i>et al.</i> ²⁰ (UK)	6 months, Lifetime	Microsimulation, Late after anti-TNF	RTX; ADA → IFX; RTX → ADA	ACR → HAQ-DI → HUI-3	ACR20/50/70	+ on treatment ++ palliative care	Direct medical
Kobelt <i>et al.</i> ²¹ (Sweden)	6 months, 10 years	Microsimulation, Early RA	ETA	HAQ-DI, DAS28 → EQ-5D	HAQ-DI, DAS28	++ after first year ++ first, second biologic + remission +++ standard therapy	Direct medical and non-medical, Indirect
Kriekkaert <i>et al.</i> ²² (the Netherlands)	3 months, 3 years	Markov cohort, NR	ADA	HAQ-DI, DAS28 → EQ-5D	EULAR/DAS28	NR	Direct medical and non-medical, Indirect
Lekander <i>et al.</i> ²³ (Sweden)	1 year, 20 years	Markov cohort, NR	ETA; ADA, ETA, IFX as a class	HAQ-DI, VAS → EQ-5D	HAQ-DI, DAS28	+ first anti-TNF ++ second anti-TNF +++ non-biologic therapy	Direct medical and non-medical, Indirect
Lekander <i>et al.</i> ³⁸ (Sweden)	1 year, 10 years	Markov cohort, NR	IFX	HAQ-DI, VAS → EQ-5D	HAQ-DI, DAS28	NR	Direct medical and non-medical, Indirect
Lindgren <i>et al.</i> ²⁴ (Sweden)	Not reported, 8 years (1999–2007)	DES, Late RA after first-line anti-TNF	Second-line anti-TNFs, RTX each → additional biologics	HAQ-DI, DAS28 → EQ-5D	NR	+ off treatment	Direct medical and non-medical, Indirect
Merkesdal <i>et al.</i> ²⁵ (Germany)	6 months, Lifetime	Microsimulation, Late RA after ETA	ADA → IFX; RTX → ADA → IFX	ACR → HAQ-DI → HUI-3	ACR20/50/70	+ on treatment ++ palliative care	Direct medical, Indirect
Nguyen <i>et al.</i> ²⁶ (US)	3 months, 5 years	Markov cohort, Moderate, severe RA	ETA; CRT; IFX; ADA; GOL	ACR → Utility (VAS)	ACR20/50/70	NR	Direct medical
Russell <i>et al.</i> ²⁷ (Canada)	6 months, 2 years	Decision tree, Late RA after sDMARDs, anti-TNFs	Different sequences of ABA, ETA, IFX	NA	DAS28	N/A	Direct medical and non-medical
Schipper <i>et al.</i> ²⁸ (the Netherlands)	3 months, 5 years	Markov cohort, Early RA	First anti-TNF → second anti-TNF → RTX	DAS28 → EQ-5D	DAS28	N/A	Direct medical and non-medical, Indirect
Soini <i>et al.</i> ²⁹ (Finland)	6 months, Lifetime	Microsimulation, Moderate, severe RA after sDMARDs	ADA, ETA, TOC each → RTX → IFX	ACR → HAQ-DI → EQ-5D	ACR20/50/70	0 bDMARDs – TOC + sDMARDs	Direct medical and non-medical
Spalding <i>et al.</i> ³⁰ (US)	1 year, Lifetime	Markov cohort, Early RA	ADA; ETA	HAQ-DI → HUI-3	HAQ-DI	+	Direct medical
Tanno <i>et al.</i> ³¹ (Japan)	6 months, Lifetime	Markov cohort, Late RA	BUC → ETA	ACR → HAQ-DI → EQ-5D	ACR20	0 ETA + sDMARDs	Direct medical

(continued)

Table 1. Continued.

Citation (Country)	Cycle length, time horizon	Model type, disease description	Biologic course of treatment, treatments evaluated*	Health state utilities calculation	Clinical effectiveness measures	Disease (HAQ-DI) progression**	Healthcare resource use and costs
Valle-Mercado <i>et al.</i> ³² (Colombia)	6 months, 10, 20 years lifetime	Markov cohort, All stages of RA	Biologic therapies as a class	HAQ-DI → EQ-5D	HAQ-DI	NR	Direct medical and non-medical, Indirect
Vera-Llonch <i>et al.</i> ³³ (US)	3 months, Lifetime, 10 years	Microsimulation, Late RA after anti-TNF	ABA	HAQ-DI → EQ-5D	HAQ-DI	+ ABA ++ MTX	Direct medical
Vera-Llonch <i>et al.</i> ³⁷ (US)	3 months, Lifetime, 10 years	Microsimulation, Late RA after MTX	ABA	HAQ-DI → EQ-5D	HAQ-DI	+ ABA ++ MTX	Direct medical
Wailoo <i>et al.</i> ³⁴ (US)	6 months, Lifetime	Microsimulation, Late RA	ADA; ETA; IFX; ANA	ACR → HAQ-DI → EQ-5D	ACR20/50	Varies by ACR response and time on treatment	Direct medical
Wu <i>et al.</i> ³⁵ (China)	6 months, Lifetime	Markov cohort, Moderate, severe RA after sDMARDs	ADA; IFX; ETA; ADA, IFX, ETA each → RTX	ACR → HAQ-DI → HUI-3	ACR20/50/70	++ palliative care	Direct medical, Indirect
Yuan <i>et al.</i> ³⁶ (US)	6 months, Lifetime	Microsimulation, Late RA	ABA; RTX	HAQ-DI → EQ-5D	HAQ-DI	+ biologics ++ MTX	Direct medical

*Biologic treatments shown here and may have been used with and/or compared with sDMARDs/palliative care.

**Relative RA disease progression, measured by the HAQ-DI, is indicated by 0 (no change), – (disease improvement), or + (disease progression); ++ (and +++) indicate that the treatments are associated with higher rate of disease progression than treatment with + (and +++). See the Supplementary Appendix for full details on these studies. ABA, abatacept; ADA, adalimumab; ANA, anakinra; BUC, bucillamine; CRT, certolizumab; DES, Discrete event simulation; GOL, golimumab; HAQ-DI, Health Assessment Questionnaire–Disability Index; HUI-3, Health Utilities Index Mark-3; IFX, infliximab; rheumatic drugs; EQ-5D, EuroQol EQ-5D; ETA, etanercept; EULAR, European League Against Rheumatism; RTX, rituximab; TNF, tumor necrosis factor; MTX, methotrexate; NA, not applicable; NR, not reported; RTX, rituximab; TNF, tumor necrosis factor.

in three out of five other dimensions: an acute phase reactant (erythrocyte sedimentation rate or C-reactive protein), patient-reported pain, physician global assessment of disease activity, patient global assessment of disease activity, and patient-reported physical functioning and disability (usually assessed by the HAQ-DI). The DAS28 measures disease activity using the number of tender and swollen joints, acute phase reactant levels, and patient-reported global health or patient-reported global assessment of disease activity. The HAQ-DI uses 20 questions to measure physical functioning and disability in eight dimensions.

Roughly the same proportion (one-third) of the models measured treatment response using either the HAQ-DI^{18,21,23,30,32,33,36–38} or the ACR criteria^{16,19,25,31}. Five models measured treatment response using the DAS28^{21,23,27,28,38}, and two models^{15,22} used the European League Against Rheumatism (EULAR) response, which is based on absolute and relative changes in the DAS28. The final model used data from the Southern Swedish Arthritis Treatment Group (SSATG) registry and from the REFLEX clinical trial to derive time-to-event data and, therefore, to track changes in clinical outcomes and response to treatment²⁴. Many models did not use the HAQ-DI to define response to treatment, despite the fact that the HAQ-DI is usually included in the ACR20/50/70 criteria and changes in HAQ-DI scores are generally reported along with the proportion of patients achieving ACR20/50/70 criteria. As a result, ACR20/50/70 response criteria were converted to HAQ-DI scores using published relationships before being used to further calculate utilities and QALYs.

Health-related quality-of-life and health state utilities

Because utility data, which are required to calculate QALYs, are not always collected during clinical trials, researchers have calculated utilities from clinical disease activity measures, such as the DAS28 or the HAQ-DI. Almost all ($n = 21$) of the 24 articles used various regression-based formulae to convert HAQ-DI scores to utilities, with almost half ($n = 10$) of these models first converting ACR20/50/70 response criteria to HAQ-DI scores based on correlations between HAQ-DI scores and ACR20/50/70 response criteria reported in key clinical trials^{16,17,19,20,25,26,29,31,34,35}. Despite the fact that different HAQ-DI-to-utilities conversion algorithms can influence utility improvements estimated by models, these differences are small³⁹. Although some of the HAQ-DI-to-utilities formulae included additional terms for age and sex, Wailoo *et al.*³⁴ used a much more comprehensive formula that included terms for baseline HAQ-DI, disease duration, and the number of previous DMARDs used.

One model converted DAS28 scores to utilities²⁸ and other models used a function that converted both HAQ-DI and DAS28 to utilities^{21,22,24}.

Disease progression

Disease progression is an important component of a RA model and should be implemented in a manner that reflects the chronic nature of the disease, the normal decline in functional ability as people age, changes in functional ability/disease states while on treatments and between treatments, and the patterns of treatment initiation/switching. Properly accounting for disease progression allows the model to appropriately capture time spent in different health states (which define QALYs accumulated) and treatment events (which define costs).

Ten models either assumed zero on-treatment disease progression or were otherwise unclear about how disease progression was handled^{17,19,22,24,26–29,31,32}. However, a number of RA researchers, including some of the co-authors of the NICE reports, have questioned the clinical validity of this assumption⁴⁰. Reflecting clinical expectations and realities, the remaining models, as well as a number of sensitivity analyses of the BRAM used by NICE^{12,13} accounted for disease progression, as measured by increases in the HAQ-DI. Half of those models assumed treatment-independent disease progression rates^{16,18,20,24,25,30,35}, while the other 10 models differentiated between progression during treatment with sDMARDs and bDMARDs^{15,17,21,23,29,31,33,34,36,37}.

Course of treatment

Models should capture the most salient aspects of treatment and disease response, but should also balance those features with the clinical realities of treatment such as the types of and sequences of therapies. Models that include courses of treatment that reflect guidelines or real-world treatment patterns have better face validity and may be more valued by HTA authorities than models of fewer therapies or simplified courses of treatment. Nine models considered treatment strategies that included sequences of more than one biologic, and only three models evaluated treatment strategies that were based on existing guidelines or that reflected the complexity of actual clinical practice^{15,16,20,24,25,27–29,32,35}. Tanno *et al.*³¹ compared two treatment strategies, one of which directly reflected the current Japanese treatment guidelines, and Kielhorn *et al.*²⁰ evaluated two treatment pathways, one of which followed NICE guidelines in effect at the time. Finckh *et al.*¹⁸ used an individual sampling method to better reflect treatment complexity and tracked patients who were treated according to three separate strategies that reflected real-world clinical complexity (pyramidal strategy, early

DMARD strategy, and an early biologic plus methotrexate strategy). The course of the disease and the treatment pathways were dependent on each patient's individual characteristics, which is not easily accomplished in a Markov cohort model, as was used by Tanno *et al.*³¹. Complex treatment sequences and treatment pathways that depend on previous treatment history/outcomes are more appropriately modeled using DES methods, which require data on the timing of events, and patient characteristics and outcomes before and after those events. In addition to the BRAM used by NICE^{12,13}, Lindgren *et al.*²⁴ also developed a DES model that used longitudinal data on 1903 patients in the SSATG from 1999–2007. Their model analyzed 8 years of rituximab use followed by up to two additional lines of biological therapy, vs three lines of biologic therapy for patients with late RA patients who failed first-line TNF-inhibitors.

Healthcare resource use and costs

In addition to QALYs, the other component of a cost-utility analysis and the resulting incremental cost-utility ratio (ICUR) is the cost of care under each treatment strategy. Although costs are usually assigned to specific healthcare resources associated with specific treatments and events, costs can also be linked to different HAQ-DI scores (and, hence, health states). Partly due to the different perspectives used by each model and different assumptions, not all models included the same set of healthcare resources (and costs). Many included a wide range of costs including all types of direct medical services (including travel in some cases)^{16–21,24,25,27,28,30,34}, while others included a more limited range such as only hospitalization- and medication-related services^{15,31,36}. Some models also included home care, adaptive aids, direct non-medical costs^{19,27,29,32,33,37}, and indirect costs^{16,18,21–25,28,32,35,38}. Data on healthcare resources use and costs have not always come from the same or consistent sources. Cost data have come from payment rates assigned to assumed healthcare resources or treatment patterns^{26,29,36}, been estimated from registry or survey data^{17,21,22,32,38}, or a combination of methods including modeling the relationship between health states and costs and expert opinion.

Recommendations for data collection

Clinical outcomes measures

Because the HAQ-DI and, to a lesser degree, the DAS28 are more commonly used than the ACR20/50/70 response criteria in clinical practice and are better correlated with RA's impact on employment and mortality risk, they can be regarded as more appropriate measures to use, at least for

the purposes of economic evaluations such as cost-utility analyses^{41–44}. We, therefore, recommend that trials and other types of studies use the HAQ-DI as the patient functioning and disability component of the ACR20/50/70 response measure and should use the DAS28 to collect information on disease activity scores. HAQ-DI and DAS28 scores should be collected at, and reported for multiple time points. We also recommend collecting HAQ-DI and DAS28 data because they are absolute measures that better capture the continuum of RA and are, therefore, more appropriate for long-term models than the ACR20/50/70 criteria, which are relative measures and less sensitive to non-linear trajectories in disease severity.

Outcomes data for different treatments may come from different sources, such as clinical trials, that may not have comparable patient populations or the level of placebo response, for example. In the absence of head-to-head trials, indirect or mixed-treatment comparison meta-analyses have emerged as an accepted and valid method recognized and recommended by HTA agencies for comparing drugs with each other using a common comparator, such as placebo^{45,46}. When studies cannot directly compare active treatments they should be designed so that their results can be combined with other published results through such meta-analytic methods.

Health-related quality-of-life and health state utilities

Trials and other studies should collect HRQoL data at multiple time points using an RA-specific instrument, such as the HAQ-DI, as well as a standard generic instrument that can be converted to utilities such as the EuroQol ED-5D⁴⁷ or the SF-6D⁴⁸, which are two validated generic instruments that have been used extensively in clinical studies to measure HRQoL and utilities. We recommend the HAQ-DI over other measures, such as the ACR20/50/70 response criteria or measures of radiographic damage¹⁸, because they typically must be converted to the HAQ-DI before being converted to utilities. This two-step process not only introduces additional uncertainty to any existing uncertainty related to the HAQ-DI-to-utilities conversion, but also requires that both relationships (ACR20/50/70-to-HAQ-DI and HAQ-DI-to-utilities) be derived from populations similar to that of the model.

The EQ-5D measures HRQoL on five dimensions (mobility, self-care, usual activities, pain and discomfort, and anxiety and depression) using two different measurement scales to describe the extent of any problems. The EQ-5D-3L version measures problems using three levels (no problems, some problems, or extreme problems) and the EQ-5D-5L version measures problems using five levels

(no problems, slight problems, moderate problems, severe problems, or extreme problems). Value sets to convert EQ-5D-3L scores to utilities are available for many countries. The SF-6D, which was derived from the larger SF-36⁴⁹ instrument, measures limitation levels (none of the time to all of the time; or none to extreme) in HRQoL on six dimensions (physical functioning, role limitations, social functioning, pain, mental health, and vitality)⁴⁹. Although the original SF-36 was not designed to directly elicit health state utilities, methods are available to convert SF-6D scores to utilities^{48,50}.

Both the EQ-5D and the SF-6D have been evaluated for use with RA patients, but with equivocal evidence about either one's superiority^{51–53}. The EQ-5D is the preferred instrument in some territories, such as the UK. However, researchers should consider including the SF-36 (SF-6D) in future RA trials based on the recent evidence on its ability to measure changes in HRQoL in response to anti-TNF therapy⁵⁴ and because the SF-6D captures more dimensions of HRQoL.

Disease progression

Trials and other studies should collect disease progression data (among all clinical measures including HAQ-DI as mentioned above) at regular intervals to provide detailed data on disease progression by treatment and stage of treatment. Longer-term follow-up studies would be helpful by providing disease progression data beyond the limited trial-based time horizons currently available, which could then be leveraged by researchers developing economic models of treatment for RA.

Course of treatment

Trials and other studies should be designed to capture actual clinical practice. When this is not possible within the context of tightly controlled randomized clinical trials or regulatory requirements, or because a trial is designed to reflect country-specific guideline care or unique patient and treatment landscapes, researchers might consider realistic or pragmatic trials or other designs that account for and record data on past therapies.

Healthcare resource use and costs

We recommend that future studies collect and report on resource use and costs in all categories of care (inpatient, outpatient, physician office visits, emergency services, surgery, medication, laboratory, therapies, and other). Furthermore, more detailed information should be reported, at least, for the most expensive categories of care (inpatient hospitalizations, surgery, medications) and for physician visits (by specialty). For example,

inpatient hospitalizations should be reported by reason for admission, surgery should be reported by specific types/reasons, and medications should be stratified by those that are and are not RA-specific. Resource use and cost data collected during a clinical study designed to reflect actual clinical practice and/or treatment guidelines can directly inform economic evaluation and modeling efforts, and strengthen the validity of the results of those models. Even the results of clinical trials not designed to reflect national treatment guidelines can be useful if they provide resource use and cost data for various treatment-related events and/or health states because those data can be used by other researchers as parameters for their economic models.

Depending on the perspective, we also recommend that future studies collect employment and lost labor productivity data using standard instruments that are reviewed elsewhere^{55–58}. In addition to standard measures of labor productivity loss, we recommend collecting salary information, especially in prospective studies and in studies that enroll individuals with and without RA.

Discussion

A carefully articulated and comprehensively described set of clinical pathways is clearly necessary, but not sufficient, to develop a valid economic model of RA; the model must also be populated with the appropriate input data to calculate valid and informative results. In this article, we have discussed five specific types of model inputs and how they are usually derived. We found that many of the recently published RA economic models used data that were not completely aligned with their corresponding concepts. In some cases, models did not fully reflect actual clinical (or guideline-recommended) care, which may have been due to a lack of necessary data. Specifically, we found that almost all of the models derived utilities using a 1- or 2-step process, used relative (ACR20/50/70) rather than absolute (HAQ-DI) measures of clinical outcomes and response to treatment, made assumptions about disease progression, evaluated stylized or simplified courses of treatment, and, in some cases, analyzed limited categories of healthcare resources and costs. Because much of these data have been, and can be, collected during clinical trials and other studies, we have made a number of recommendations for the design of these studies including how these data can be collected. These data can, in turn, be used by researchers developing economic models and conducting economic evaluations of new RA treatments. Our recommendations are meant to provide guidance to the design of future clinical and other studies so that the most appropriate outcomes and economic data can be collected to support economic analyses.

The validity of CUA models depends on how well results from short-term clinical trials can be extrapolated over longer time horizons. Therefore, in addition to our specific recommendations to collect data on HRQoL and utilities, relevant clinical outcome, disease progression, course of treatment, and healthcare resource use and costs, we also recommend that future studies collect these and other data, such as adverse events, over as long as possible time periods and at multiple time points to fully capture periods of remission and relapse and at different levels of physical functioning and disability.

Our analyses and recommendation are not exhaustive. Key limitations of our approach include the relatively small number of model characteristics we examined and our literature review selection criteria, which may have excluded relevant models or other types of evaluations of bDMARDs from review. We also recognize that not all of these recommendations may be feasible in the immediate future and that not all clinical trials and other studies will have the resources to support long-term follow-up periods. However, to the extent that these recommendations can be implemented, future models that use data from these enhanced data collection efforts may have results with higher degrees of validity and credibility.

Conclusions

A well-designed economic model is necessary, but not sufficient to validly and informatively evaluate and compare competing biologic treatments for RA. Models must also be populated with appropriate data that capture all aspects of the disease, its treatments, and its outcomes. We recommend that future studies collect clinical outcomes and health-related quality-of-life data using appropriate instruments that can convert directly to utilities; collect data on actual disease progression; be designed to capture real-world courses of treatment; and collect detailed data on a wide range of healthcare resources and costs.

Transparency

Declaration of funding

This study was funded by Novo Nordisk A/S.

Declaration of financial/other relationships

MLG is an employee of Evidera, formerly a division of United BioSource Corporation. United BioSource Corporation received funding from Novo Nordisk A/S for conducting a literature review as a basis for this study. BBH, MS-L, and XV were all employees at Novo Nordisk A/S or Novo Nordisk Inc. while this work was performed. XV is currently at Bristol-Myers Squibb. Decisions regarding study design, interpretation, and writing and submitting the manuscript were jointly made by the authors.

Acknowledgments

The authors would like to thank Meghan Burns and Amber Martin for their assistance preparing this manuscript.

References

- Singh JA, Furst DE, Bharat A, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2012;64:625-39
- Smolen JS, Landewe R, Breedveld FC, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. *Ann Rheum Dis* 2014;73:492-509
- Marra CA, Bansback N, Anis AH, et al. Introduction to economic modeling for clinical rheumatologists: application to biologic agents in rheumatoid arthritis. *Clin Rheumatol* 2011;30(Suppl 1):S9-18
- Gabriel S, Drummond M, Maetzel A, et al. OMERACT 6 Economics Working Group report: a proposal for a reference case for economic evaluation in rheumatoid arthritis. *J Rheumatol* 2003;30:886-90
- Maetzel A, Tugwell P, Boers M, et al. Economic evaluation of programs or interventions in the management of rheumatoid arthritis: defining a consensus-based reference case. *J Rheumatol* 2003;30:891-6
- Drummond MF, Barbieri M, Wong JB. Analytic choices in economic models of treatments for rheumatoid arthritis: What makes a difference? *Med Decis Making* 2005;25:520-33
- Barbieri M, Wong JB, Drummond M. The cost effectiveness of infliximab for severe treatment-resistant rheumatoid arthritis in the UK. *Pharmacoeconomics* 2005;23:607-18
- Barton P, Jobanputra P, Wilson J, et al. 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
- Kobelt G, Jonsson L, Young A, et al. 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;42:326-35
- National Institute for Health and Care Excellence. Guidance on the Use of Etanercept and Infliximab for the Treatment of Rheumatoid Arthritis. London: National Institute for Health and Care Excellence; 2002
- National Institute for Health and Care Excellence. The Clinical Effectiveness and Cost-Effectiveness of New Drug Treatments for Rheumatoid Arthritis: Etanercept and Infliximab. Technology Appraisal No. 36. London: National Institute for Health and Care Excellence; 2002
- Chen YF, Jobanputra P, Barton P, 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;10:1-229
- Malottki K, Barton P, Tsourapas A, 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;15:1-300
- Beresniak A, Russell AS, Haraoui B, et al. Advantages and limitations of utility assessment methods in rheumatoid arthritis. *J Rheumatol* 2007;34:2193-200
- Brennan A, Bansback N, Nixon R, 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;46:1345-54
- Davies A, Cifaldi MA, Segurado OG, et al. Cost-effectiveness of sequential therapy with tumor necrosis factor antagonists in early rheumatoid arthritis. *J Rheumatol* 2009;36:16-26
- Diamantopoulos A, Benucci M, Capri S, et al. Economic evaluation of tocilizumab combination in the treatment of moderate-to-severe rheumatoid arthritis in Italy. *J Med Econ* 2012;15:576-85
- Finckh A, Bansback N, Marra CA, 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;151:612-21
- Hallinen TA, Soini EJ, Eklund K, et al. Cost-utility of different treatment strategies after the failure of tumour necrosis factor inhibitor in rheumatoid arthritis in the Finnish setting. *Rheumatology (Oxford)* 2010;49:767-77
- Kielhorn A, Porter D, Diamantopoulos A, et al. 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;24:2639-50
- Kobelt G, Lekander I, Lang A, et al. Cost-effectiveness of etanercept treatment in early active rheumatoid arthritis followed by dose adjustment. *Int J Technol Assess Health Care* 2011;27:193-200
- Krieckaert CLM, Nair SC, Nurmoahamed MT, et al. Personalised treatment using serum drug levels of adalimumab in patients with rheumatoid arthritis: An evaluation of costs and effects. *Ann Rheum Dis* 2013;0:1-8. doi:10.1136/annrheumdis-2013-204101
- Lekander I, Borgstrom F, Lysholm J, et al. The cost-effectiveness of TNF-inhibitors for the treatment of rheumatoid arthritis in Swedish clinical practice. *Eur J Health Econ* 2013;14:863-73
- 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;25:181-9
- Merkesdal S, Kirchhoff T, Wolka D, et al. Cost-effectiveness analysis of rituximab treatment in patients in Germany with rheumatoid arthritis after etanercept-failure. *Eur J Health Econ* 2010;11:95-104
- Nguyen CM, Bounthavong M, Mendes MA, 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;30:575-93
- Russell A, Beresniak A, Bessette L, 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;28:403-12
- Schipper LG, Kievit W, den Broeder AA, et al. Treatment strategies aiming at remission in early rheumatoid arthritis patients: Starting with methotrexate monotherapy is cost-effective. *Rheumatology (Oxford)* 2011;50:1320-30
- Soini EJ, Hallinen TA, Puolakkka K, et al. Cost-effectiveness of adalimumab, etanercept, and tocilizumab as first-line treatments for moderate-to-severe rheumatoid arthritis. *J Med Econ* 2012;15:340-51
- Spalding JR, Hay J. Cost effectiveness of tumour necrosis factor-alpha inhibitors as first-line agents in rheumatoid arthritis. *Pharmacoeconomics* 2006;24:1221-32
- Tanno M, Nakamura I, Ito K, et al. Modeling and cost-effectiveness analysis of etanercept in adults with rheumatoid arthritis in Japan: a preliminary analysis. *Mod Rheumatol* 2006;16:77-84
- Valle-Mercado C, Cubides MF, Parra-Torrido M, et al. Cost-effectiveness of biological therapy compared with methotrexate in the treatment for rheumatoid arthritis in Colombia. *Rheumatol Int* 2013;33:2993-7
- Vera-Llonch M, Massarotti E, Wolfe F, 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;35:1745-53
- Wailoo AJ, Bansback N, Brennan A, et al. Biologic drugs for rheumatoid arthritis in the Medicare program: a cost-effectiveness analysis. *Arthritis Rheum* 2008;58:939-46
- Wu B, Wilson A, Wang FF, et al. Cost effectiveness of different treatment strategies in the treatment of patients with moderate to severe rheumatoid arthritis in China. *PLoS One* 2012;7:e47373
- Yuan Y, Trivedi D, Maclean R, et al. Indirect cost-effectiveness analyses of abatacept and rituximab in patients with moderate-to-severe rheumatoid arthritis in the United States. *J Med Econ* 2010;13:33-41

37. Vera-Llonch M, Massarotti E, Wolfe F, et al. Cost-effectiveness of abatacept in patients with moderately to severely active rheumatoid arthritis and inadequate response to methotrexate. *Rheumatology (Oxford)* 2008;47:535-41
38. Lekander I, Kobelt G, Svarvar P, et al. The comparison of trial data-based and registry data-based cost-effectiveness of infliximab treatment for rheumatoid arthritis in Sweden using a modeling approach. *Value Health* 2013;16:251-8
39. Strandberg-Larsen M, Hansen BB, Göthberg M, et al. Outcome differences in algorithms used for indirect mapping of utility values from HAQ-DI: An assessment based on Phase 2a clinical trial data in patients with rheumatoid arthritis after treatment With NNC0109-0012 (Anti-IL-20 MAB). *Value Health* 2012;15:A476 [Abstract No. PRM490]
40. Jobanputra P. A clinician's critique of rheumatoid arthritis health economic models. *Rheumatology (Oxford)* 2011;50(Suppl 4):iv48-iv52
41. Bansback N, Marra C, Tsuchiya A, et al. Using the health assessment questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis. *Arthritis Rheum* 2007;57:963-71
42. Farragher TM, Lunt M, Bunn DK, et al. Early functional disability predicts both all-cause and cardiovascular mortality in people with inflammatory polyarthritis: results from the Norfolk Arthritis Register. *Ann Rheum Dis* 2007;66:486-92
43. Kessler RC, Maclean JR, Petukhova M, et al. The effects of rheumatoid arthritis on labor force participation, work performance, and healthcare costs in two workplace samples. *J Occup Environ Med* 2008;50:88-98
44. Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Soc Sci Med* 2005;60:1571-82
45. Schoels M, Aletaha D, Smolen JS, et al. Comparative effectiveness and safety of biological treatment options after tumour necrosis factor alpha inhibitor failure in rheumatoid arthritis: systematic review and indirect pairwise meta-analysis. *Ann Rheum Dis* 2012;71:1303-8
46. Spinner DS, Birt J, Walter JW, et al. Do different clinical evidence bases lead to discordant health-technology assessment decisions? An in-depth case series across three jurisdictions. *Clinicoecon Outcomes Res* 2013;5:69-85
47. The EuroQol Group. EuroQol—A new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199-208
48. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271-92
49. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83
50. Kharroubi SA, Brazier JE, Roberts J, et al. Modelling SF-6D health state preference data using a nonparametric Bayesian method. *J Health Econ* 2007;26:597-612
51. Adams R, Walsh C, Veale D, et al. Understanding the relationship between the EQ-5D, SF-6D, HAQ and disease activity in inflammatory arthritis. *Pharmacoeconomics* 2010;28:477-87
52. Salaffi F, Carotti M, Ciapetti A, et al. A comparison of utility measurement using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2011;29:661-71
53. Sorensen J, Linde L, Ostergaard M, et al. Quality-adjusted life expectancies in patients with rheumatoid arthritis—comparison of index scores from EQ-5D, 15D, and SF-6D. *Value Health* 2012;15:334-9
54. Buitinga L, Braakman-Jansen LM, Taal E, et al. Comparative responsiveness of the EuroQol-5D and Short Form 6D to improvement in patients with rheumatoid arthritis treated with tumor necrosis factor blockers: results of the Dutch Rheumatoid Arthritis Monitoring registry. *Arthritis Care Res (Hoboken)* 2012;64:826-32
55. Beaton DE, Tang K, Gignac MA, et al. Reliability, validity, and responsiveness of five at-work productivity measures in patients with rheumatoid arthritis or osteoarthritis. *Arthritis Care Res (Hoboken)* 2010;62:28-37
56. Burton W, Morrison A, Maclean R, et al. Systematic review of studies of productivity loss due to rheumatoid arthritis. *Occup Med (Lond)* 2006;56:18-27
57. Escorpizo R, Bombardier C, Boonen A, et al. Worker productivity outcome measures in arthritis. *J Rheumatol* 2007;34:1372-80
58. Filipovic I, Walker D, Forster F, et al. Quantifying the economic burden of productivity loss in rheumatoid arthritis. *Rheumatology (Oxford)* 2011;50:1083-90

Appendix

Medline search strategy

Search #	Search terms
1	'arthritis, rheumatoid'[Mesh] OR 'rheumatoid arthritis'[tiab] OR ('rheumatoid'[tiab] AND 'arthritis'[tiab])
2	cost*[tiab] OR economic*[tiab] OR 'cost-effectiveness'[tiab] OR (cost[tiab] AND effectiveness[tiab]) OR 'cost-utility'[tiab] OR (cost[tiab] AND utility[tiab]) OR pharmacoeconomic*[tiab] OR 'decision analysis'[tiab] OR 'decision analytic'[tiab] OR (decision[tiab] AND analy*[tiab]) OR markov[tiab] OR simulation[tiab] OR 'decision tree'[tiab] OR 'decision model'[tiab]
3	#1 AND #2
4	#3 Limits: Humans, Publication Date from 2006/01/01 to 2014/02/18

Embase search strategy

Search #	Search terms
1	'arthritis, rheumatoid'/exp OR 'rheumatoid arthritis':ti,ab OR (rheumatoid:ti,ab AND arthritis:ti,ab)
2	cost*:ti,ab OR economic*:ti,ab OR 'cost-effectiveness':ti,ab OR (cost:ti,ab AND effectiveness:ti,ab) OR 'cost-utility':ti,ab OR (cost:ti,ab AND utility:ti,ab) OR pharmacoeconomic*:ti,ab OR 'decision analysis':ti,ab OR 'decision analytic':ti,ab OR markov:ti,ab OR simulation:ti,ab OR 'decision tree':ti,ab OR 'decision model':ti,ab
3	#1 AND #2
4	#3 Limits: Humans, Publication Date from 2006/01/01 to 2014/02/18

Cochrane search strategy

Search #	Search terms
1	((rheumatoid arthritis):ti,ab,kw) AND (rheumatoid:ti,ab,kw OR arthritis:ti,ab,kw)
2	cost*:ti,ab OR economic*:ti,ab OR ((cost-effectiveness):ti,ab) OR (cost:ti,ab AND effectiveness:ti,ab) OR ((cost-utility):ti,ab) OR (cost:ti,ab AND utility:ti,ab) OR pharmacoeconomic*:ti,ab OR ((decision analysis):ti,ab) OR ((decision analytic):ti,ab) OR (decision:ti,ab AND analy*:ti,ab) OR markov:ti,ab OR simulation:ti,ab OR ((decision tree):ti,ab) OR ((decision model):ti,ab)
3	#1 AND #2
4*	(#3), from 2006 to 2014/02/18