The Sheffield rheumatoid arthritis health economic model

Jonathan Tosh¹, Alan Brennan¹, Allan Wailoo¹ and Nick Bansback²

Abstract

The Sheffield RA health economic model has been used in several published cost-effectiveness analyses in both the UK and internationally to evaluate different treatments for patients with RA. This article presents the key methods and assumptions that underpin the model, including justifications for using an individual patient sampling methodology, and why the model has used the HAQ to track disease activity. The article also details how trial and observational data are used in the model to address specific questions. The model has been used to support health policy in both the UK and internationally, although the limited evidence still provides a challenge when using an economic model to determine the cost-effectiveness of RA treatments. The results of analyses using the Sheffield RA model are presented. The limitations of the model are discussed, and improvements are continually required to provide a model that is appropriate to address health economic questions in the future. The Sheffield RA model continues to be used and refined, and allows health economic questions to be answered using a transparent and flexible modelling methodology.

Key words: Rheumatoid arthritis, Economic, Cost, Effectiveness, Sheffield, Model, DMARDs, Biologics, Simulation.

Introduction

The development of biological DMARDs (bDMARDs) has seen a revolution in the treatment of patients with RA. Traditionally, patients were treated with analgesics, anti-inflammatories, steroids and DMARDs. These medical treatments are relatively inexpensive and are able to reduce disease activity, but are unable to halt erosive damage and reduced functional damage caused by the progressive nature of the disease. Treatment of RA typically involves use of different therapies over the long term, with some patients having long periods of successful improvement in symptoms, while others withdraw from them more quickly due to lack of efficacy or adverse events and move onto an alternative. Biologic therapies, including TNF-α inhibitors, are better able to reduce the level of the disease, as well as slow or halt the damage to patients’ joints. However, biologics are also considerably more expensive than traditional therapies and represent a greater burden on limited health resources.

The constraints on health-care resources mandate that resource-allocation decisions be guided by considerations of the cost in relation to expected benefits [1]. A cost-effectiveness analysis is a method of economic evaluation that considers both the additional health benefits and the additional costs of a new treatment to determine whether they represent value for money. Health benefits can be measured in terms of quality-adjusted life years (QALYs), a single measure that combines length of life with health-related quality of life (HRQoL) [2]. This is achieved by assigning every health state an HRQoL value, where 1 is equal to full health and states as bad as dead are given 0. A chronic health state lasting 10 years with an HRQoL of 0.3 equates to three QALYs. Commonly used HRQoL instruments include the EQ-5D and the short-form 36 health survey (SF-36), which are generic measures of HRQoL. Disease specific measures are also available, such as the European Organisation for Research and Treatment of Cancer instrument (for cancer) and the Multiple Sclerosis Quality of Life-54 instrument (for multiple sclerosis); however, as yet, there is no RA-specific HRQoL measure. It is generally more realistic that patients experience many different health states throughout their lives, and so it is common for analyses to categorize the individual’s experience into different health states with HRQoL scores assigned. Lifetime QALYs are calculated by summing the QALYs for each health state.
To compare alternative treatments, incremental costs and incremental QALYs between treatments are calculated. The ratio of incremental costs and incremental QALYs is called the incremental cost effectiveness ratio (ICER), which is the key metric estimated in a cost-effectiveness analysis. The National Institute for Health and Clinical Excellence (NICE) is an independent organization in the UK that provides guidance on the use of treatments in the NHS. Cost effectiveness analyses are a key component of the NICE appraisal process of new treatments, and NICE has a threshold range of £20 000–£30 000 per additional QALY. If treatments have an ICER above this threshold, it is unlikely they will get a positive recommendation. Due to the chronic nature of RA, NICE requires a cost-effectiveness analysis that estimates the costs and benefits of a treatment that accrue over a patient’s remaining lifetime [3]. Decision-analytic modelling methods are required to extrapolate short-term clinical trial evidence, as well as incorporating longer term observational data [4].

The School of Health and Related Research (ScHARR) at the University of Sheffield has undertaken several cost-effectiveness analyses of treatments for RA. The analyses have been undertaken using an individual patient simulation (IPS) model constructed in Microsoft Excel (referred to now as the Sheffield RA model). This article will explain the key methods and assumptions that underpin the Sheffield RA model. The article will then briefly explain how the model has been used in various published cost-effectiveness analyses. Finally, the article will discuss how decision-analytic models could be improved when undertaking cost-effectiveness analyses of RA treatments in the future.

The Sheffield RA model

The Sheffield RA model uses IPS methods to estimate the total costs and QALYs of each treatment strategy. An IPS model generates a simulated patient with a set of baseline characteristics; in this case their age, gender, HAQ score, disease duration and previous therapies. The benefit of an IPS model over more commonly used methods, such as cohort decision trees and Markov models, is that the patients’ future in the model is conditional on their past. For example, the model may allow progression to certain treatments if a specific condition is met (i.e. number of previous treatments or a certain disease level). Cohort models can be used to incorporate different patient subgroups, but very quickly become unmanageable and inefficient if they are defined to any degree of complexity [5]. Decision trees are said to become decision bushes with huge amounts of branches or unique health states. In an IPS model, incorporating patient’s attributes and history is limited only by the availability of relevant data. There are numerous potential variables that are relevant to patient response and progression, but there is not always the data available to incorporate them into the model.

The IPS method also allows different patient subpopulations (i.e. early RA, DMARD experienced) to be evaluated in the model. The model evaluates the patient’s HAQ score at 6-monthly intervals for their whole lifetime. This allows a full sequence of treatments to be evaluated. It is the sequential use of therapies over time and the uncertain duration of effect on each patient that makes RA a suitable area for using IPS methods. The model is typically run with a large number of simulated patients to estimate the mean total cost and mean total QALYs.

HAQ

The model uses the HAQ as a proxy for patient’s disease activity. It is recognized that the HAQ is a functional measure, and does not capture the full impact of RA on patient’s quality of life. However, evaluating a patient’s HAQ allows a conversion to a HRQoL outcome using any of a number of published linear regression functions [6]. These functions have demonstrated a relatively strong correlation between the HAQ and several HRQoL instruments. While this method is a widely used way to bypass the problem posed due to many trials not measuring HRQoL, it is not without controversy. Some versions of the Sheffield RA model have used observational studies that have measured both HRQoL and HAQ to estimate HRQoL directly.

Figure 1 shows a potential HAQ profile for a patient simulated in the model. The patient has a simulated baseline HAQ, and after receiving the first treatment sees an improvement (decrease) in their HAQ score. While on treatment, their HAQ level worsens over time, until a point comes when the treatment is stopped and their HAQ rebounds. In Fig. 1, the patient begins a second therapy that is less effective than the first, both in terms of the response gained and the time spent on treatment.

The Sheffield RA model requires three key pieces of data: the initial effectiveness of a treatment; the change over time in a patient’s HAQ while on treatment; and the length of time a patient will spend on treatment. Typically, randomized controlled trial data have been used in the Sheffield RA model to provide the first set of data, while observational studies have been used to provide the longer run data. Data-informed assumptions are required regarding how a patient’s HAQ will rebound after treatment is withdrawn. In the analyses undertaken, patients are assumed to withdraw either due to adverse events or to a loss of efficacy. In practice, this phenomenon is when a patient’s disease flares and disease control is lost. Adverse events will see a patient withdrawn from a treatment and progress to the next in the sequence. The model does not explicitly model the costs and health effects of adverse events. The model allows a number of different possibilities regarding what happens to a patient’s HAQ after treatment withdrawal:

(i) a rebound that is equal to the initial gain when a treatment is initiated;
(ii) a rebound back to a patient’s baseline HAQ; and
(iii) a rebound to the point that corresponds with a patient’s potential disease level if no treatment was introduced.

This assumption is crucial for any analysis, because it first implies any assumption about a treatment’s ability to
delay or halt a patient’s disease progression, and whether this effect lasts after treatment withdrawal. Also, differential assumptions between treatments within an analysis can be a key driver for the cost-effectiveness results.

Modelling methods

The model estimates costs and resource use in two ways. First, direct treatment costs are estimated, and secondly the model allows the use of hospitalizations and resources used to be linked to the HAQ profile of the patient. Either observational data or published analyses provide an estimate of the number of hospitalizations, surgical interventions and joint replacements as a function of a patient’s HAQ level. This allows resources used due to disease activity, functional damage and serious adverse events to be appropriately costed without requiring an extra level of complexity to be explicitly modelled.

The model is probabilistic, in that it incorporates the uncertainty around parameter inputs and uses Monte Carlo sampling methods to quantify the uncertainty. This first accounts for the fact that input parameters may be skewed and models are unlikely to be linear, and therefore allows an accurate estimate of mean costs and QALYs. Quantifying uncertainty also identifies areas of uncertainty or sensitivity, which therefore drives future research and will improve future decision making.

Etanercept analysis

The Sheffield RA model was first used to evaluate etanercept as a third-line treatment compared with sequential DMARD therapy [7]. The analysis was performed for Wyeth, as part of their submission to the NICE Technology Appraisal process. The analysis provided an estimated ICER of £16,330 per additional QALY for etanercept compared with sequential DMARDs. The uncertainty around the inputs into the model was quantified in a sensitivity analysis, which provided a plausible range of ICER estimates of £7,800–£42,000 per additional QALY. The sensitivity analysis found that long-term HAQ progression rates for etanercept, DMARDs and non-responding patients were the most sensitive variables.

As part of the NICE Technology Appraisal process, an independent academic group was commissioned to develop a separate model, and this analysis estimated an ICER of between £27,000 and £35,000 per additional QALY. The evidence presented by Wyeth and by the independent academic group was considered by the NICE appraisal committee before etanercept was given its positive recommendation.

British Society of Rheumatology Biologics Registry analysis

In 2006, we collaborated with the British Society of Rheumatology Biologics Registry (BSRBR) to undertake an independent evaluation of TNF-α inhibitors as third-line treatments compared with sequential DMARD monotherapy [8]. The BSRBR recruited all RA patients starting TNF-α inhibitors in the UK since 2001, and collected 6-monthly data including the HAQ and DAS-28 clinical measures as well as the SF-36 and EQ-5D HRQoL instruments. The BSRBR data set was used to derive estimates of short-run efficacy, the time spent on treatment, as well as the costs of treatment, monitoring and hospitalizations. The CEA analysis reported an ICER of £23,888 per additional QALY for TNF-α inhibitors (as a group) compared with sequential DMARDs. A sensitivity analysis was performed which found that the results were sensitive to the withdrawal rules, the discount rates used and the HAQ progression rates. The probabilistic sensitivity analysis (PSA) showed that there was an 84% probability of TNF-α inhibitors being cost-effective at the £30,000 per QALY threshold.

Medicare analysis

In 2008, ScHARR was commissioned by the US Agency for Healthcare Research and Quality (AHRQ) to undertake an analysis of bDMARDs in bDMARD-naïve patients [9]. This analysis was performed as a part of the Medicare
Modernization Act. The analysis incorporated data from the National Databank for Rheumatic Diseases (NDB), a US RA registry. At the time of the analysis, the registry included over 3000 patients on TNF-α inhibitors. The cost-effectiveness analysis suggests recommending the use of etanercept or adalimumab, and not infliximab or anakinra. Anakinra was both the least costly and least effective strategy. Etanercept, adalimumab and infliximab were all similar in terms of effectiveness; however, infliximab is more costly and therefore not likely to be cost-effective. The basecase results are presented in Table 1. Medicare does not have a cost-effectiveness threshold, but the analysis determined that if decision makers are willing to pay a maximum of $50,000 per additional QALY, then the probability of infliximab being cost-effective is <1%. The sensitivity analysis performed found the results were generally robust, although if no dose increase is assumed for infliximab then it becomes a cost-effective strategy compared with the other TNF-α inhibitors.

**DMARD analysis**

The Sheffield RA model was further developed and used during the NICE Clinical Guideline for Rheumatoid Arthritis [10–12]. SCARR was commissioned to undertake the health economics component of the guideline development. This analysis focused on early RA and evaluated different combinations of DMARD strategies in patients with early RA. The objective was to determine the appropriate treatment with DMARDs before patient’s progress to TNF-α inhibitors as determined by existing NICE guidance. Five unique combination DMARD strategies were compared, along with sequential DMARD monotherapy. Combinations of DMARDs that either titrate dosage downwards, or involve intensive triple DMARDs are likely to be cost-effective compared with monotherapy. At a £20,000 per QALY threshold, then the strategies most likely to be cost-effective were either DMARD combination therapy with downwards titration (probability of being optimal = 50%) or intensive, triple DMARD therapy (probability of being optimal = 43%). The results were robust to a range of scenario analyses.

**Discussion**

The Sheffield RA model has shown its flexibility, both in terms of evaluating different populations of patients, and being able to utilize a wide range of data types. For instance, in the AHRQ analysis, the large NDB data set allowed an analysis of a wide range of covariates. This flexibility has future-proofed the model, and will allow adaptations to address other health-economic questions in RA. Since the model can evaluate sequences, it has been relatively simple to adjust so to reflect changing clinical practice.

The Sheffield RA model is one of several RA models developed in recent years [13, 14]. While the presence of multiple models may be confusing at first, particularly when they seemingly give different results, it can be important to compare analyses using different assumptions and frameworks to ensure the validity of results. Models can produce different results for two reasons: due to alternative model structures, and due to different model inputs. Models must be understood by their target audience and trusted by the end users. Putting similar inputs into different models to formally appraise models, as has been done in diabetes and cancer would be beneficial [15, 16]. Such issues are not confined to economic models; a recent exercise in which researchers developed seven independent statistical models of breast cancer incidence and mortality to assess the effects of screening and adjuvant therapy reported substantial variability across models [17].

The primary difference between model structures of previous developed analyses is the use of IPS in contrast to a cohort framework. The reason for using an IPS model framework has been to allow conditional rules to be modelled. RA has treatment decisions based on clinical rules, including the DAS-28 5.1 score requirement for progression to biologics. IPS allows these rules to be both implemented and evaluated. Our opinion is also that the Sheffield RA model has a key advantage in that it explicitly follows treatment success or failure as determined in clinical practice, and failures are withdrawn and the costs and benefits of responders are followed into the longer term.

However, IPS modelling has been criticized by some as a less transparent method, and some have not incorporated probabilistic sensitivity analysis due to the computational time required to perform [18]. The use of IPS has grown in recent years, and it is hoped that this article will have provided more detail on why IPS modelling is required in RA, as well as providing a transparent explanation of the methods of the Sheffield RA model.

The key issue as always when developing a model is the limited evidence base. RA has numerous clinical trials; however, they often only compare a treatment with

**Table 1** Medicare analysis base case results

<table>
<thead>
<tr>
<th></th>
<th>Mean cost, $</th>
<th>Mean QALYs</th>
<th>ICER (compared with infliximab)</th>
<th>Incremental ICER</th>
<th>Mean treatment duration, years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infliximab</td>
<td>94029</td>
<td>7.64</td>
<td>Dominates</td>
<td>Dominated</td>
<td>4.67</td>
</tr>
<tr>
<td>Etanercept</td>
<td>81181</td>
<td>7.66</td>
<td>Dominates</td>
<td>92058 compared with adalimumab</td>
<td>5.00</td>
</tr>
<tr>
<td>Adalimumab</td>
<td>79535</td>
<td>7.64</td>
<td>Dominates</td>
<td>142726 compared with anakinra</td>
<td>4.97</td>
</tr>
<tr>
<td>Anakinra</td>
<td>50608</td>
<td>7.44</td>
<td></td>
<td>–</td>
<td>1.76</td>
</tr>
</tbody>
</table>
placebo, rather than against all relevant comparator treatments. Therefore, network meta-analysis techniques are becoming necessary to estimate the comparative effectiveness of different therapies. These methods have proven crucial for numerous cost-effectiveness analyses; however, they require transparent reporting regarding the appropriate statistical model selected, and also require the evidence base to provide a closed network of treatments for comparison [19, 20]. Because short-term trials are unable to provide the longer term data required for the model, observational data are required which, due to their non-randomized nature, may provide biased estimates of treatment effectiveness. Also, observational data, by their nature are historical, and the care and service provided at the time may have evolved.

The analyses described in this article are only a sub-set of the analyses and publications based upon the Sheffield RA model, including an evaluation of adalimumab for use in Sweden [21]. The model has other uses, such as in clinical trial development [22], and Value of Information and PSA methods have been performed using the model, which have informed the ongoing RA research agenda.

Conclusions

The Sheffield RA model has provided a flexible basis for economic evaluations of RA therapies over the last few years. Decisions and recommendations concerning the use of bDMARDs in the UK, Sweden and USA, and combination strategies for the use of traditional DMARDs in early RA, have been supported by the use of a consistent health economic framework. The model has been able to incorporate and synthesis new data from the key clinical trials and from longer term observational studies in the developing evidence base. Validation of model structure, assumptions and evidence are required to provide a transparent cost-effectiveness analysis. This continues to be the case for a complex disease like RA, where models need to incorporate evidence from a range of sources, and where evidence is limited for parameters that are particularly sensitive.

Supplement: This paper forms part of the supplement ‘Biologic therapies in inflammatory joint diseases: models, evidence and decision making’. This supplement was supported by unrestricted funding from Arthritis Research UK.

Disclosure statement: The authors have declared no conflicts of interest.

References


