Development of the Birmingham Rheumatoid Arthritis Model: past, present and future plans

Pelham Barton¹

Abstract

The Birmingham Rheumatoid Arthritis Model (BRAM) has been developed over a number of years to inform several appraisals of biologic drugs by the Technology Appraisals Committee of the UK National Institute for Health and Clinical Excellence. This article describes the processes used in the construction of the different versions of the BRAM.

Key words: Rheumatoid arthritis, NICE, Modelling, Biologic therapies, Cost-effectiveness analysis.

Introduction

The National Institute for Health and Clinical Excellence (NICE) includes in its functions the technology appraisals programme, which in practice is mostly concerned with the appraisal of therapeutic drugs. Technologies can now be assessed through two different processes known as Multiple Technology Appraisal and Single Technology Appraisal [1]. In the Multiple Technology Appraisal process, an independent team (known as the assessment group) is commissioned to produce an assessment report including an economic model. The West Midlands Health Technology Assessment Collaboration was commissioned to produce such a report for four separate Multiple Technology Appraisals involving biologic drugs for the treatment of RA. This article describes the development of the modelling used in those four reports.

The Birmingham preliminary model

The first appraisal of biologic drugs for RA was for etanercept and infliximab. Our assessment report [2] formed part of the evidence considered in the development of the final guidance [3]. The basic question addressed in that appraisal was whether etanercept and infliximab should be available for routine use in the National Health Service (NHS) within their licensed indications. To perform this assessment, we took as the comparator a specific sequence of non-biologic DMARDs. This sequence (Fig. 1), beginning with SSZ, MTX and gold (sodium aurothiomalate), was chosen to represent a typical pattern that might be used in the absence of biologic therapy. It was assumed that any patient who survived sufficiently long would quit any treatment for reasons of either toxicity or loss of effectiveness, and that any patient who survived the whole sequence of active treatments would move on to palliation. We then considered the effect of including either etanercept or infliximab in two possible places in the sequence: as third-line therapy, after MTX and before gold, or as the last active treatment, after all traditional DMARDs and before palliation.

Structure of the model

Information available to us suggested that the time to quitting any treatment did not follow an exponential distribution, and we opted for a continuous distribution of time to quitting. This choice meant that the appropriate choice of model was the type known as an individual sampling model [4]. In such a model, a sequence of virtual patient histories is generated with uncertain outcomes determined by sampling using a computer random number generator. We built the model in TreeAge software (TreeAge Software Inc; www.treeage.com).

Each virtual patient history is constructed as follows. First, a set of initial patient characteristics is sampled from an appropriate distribution, usually multivariate. In fact in the Birmingham preliminary model, the only patient characteristic required was the patient’s remaining lifetime. To perform this assessment, we took as the comparator a specific sequence of non-biologic DMARDs. This sequence (Fig. 1), beginning with SSZ, MTX and gold (sodium aurothiomalate), was chosen to represent a typical pattern that might be used in the absence of biologic therapy. It was assumed that any patient who survived sufficiently long would quit any treatment for reasons of either toxicity or loss of effectiveness, and that any patient who survived the whole sequence of active treatments would move on to palliation. We then considered the effect of including either etanercept or infliximab in two possible places in the sequence: as third-line therapy, after MTX and before gold, or as the last active treatment, after all traditional DMARDs and before palliation.

¹Health Economics Unit, School of Health and Population Sciences, University of Birmingham, Birmingham, UK.

Submitted 2 February 2011; revised version accepted 7 June 2011.

Correspondence to: Pelham Barton, Health Economics Unit, School of Health and Population Sciences, Public Health Building, University of Birmingham, Birmingham B15 2TT, UK.

E-mail: p.m.barton@bham.ac.uk

The Author 2011. Published by Oxford University Press on behalf of the British Society for Rheumatology. All rights reserved. For Permissions, please email: journals.permissions@oup.com
In the Birmingham preliminary model, we assumed that the effect of any treatment was to produce a fixed improvement in the patient’s health-related quality of life, and that this improvement would be lost on quitting treatment (Fig. 2). This assumption allowed QALYs to be modelled relative to the basic curve (Fig. 3). We made no specific assumption about the underlying pattern of change in quality of life over the patient’s lifetime.

Costs included in the model were treatment costs and monitoring costs, including administration costs for infliximab and appropriate general practitioner and hospital visits. Pretreatment tests were included as a one-off cost on starting treatment. For simplicity, additional costs of heavier monitoring or higher doses early in the course of a particular treatment were included in this starting cost.

QALYs were counted relative to the assumed (and unspecified) pattern of background deterioration in quality of life. For each active treatment, a fixed improvement was applied for the duration of that treatment. Deductions were made at the start and end of each treatment representing, respectively, the delay in a treatment taking effect and an assumed gradual loss of benefit before switching treatment (either loss of effectiveness or onset of toxicity).

In line with the requirements of NICE at the time, costs and QALYs were discounted at 6 and 1.5%, respectively. Results were presented using both the starting point of the model and the point of divergence between strategies as the base point for discounting; the latter is the technically correct approach.

Limitations of the Birmingham preliminary model
We drew attention in our assessment report [2] to a number of significant limitations within the Birmingham preliminary model. These included the fact that no severity-related mortality effects could be modelled, and the model took no account of joint replacement or hospitalization.

The Birmingham rheumatoid arthritis model (first version)
Following the development of the Birmingham preliminary model, members of the West Midlands Health Technology Assessment Collaboration knew that we were likely to be the assessment group for further appraisals of biologic therapies for RA. Accordingly, we arranged with the National Co-ordinating Committee for Health Technology Assessment to use part of our allocation of research time within our contract to develop the model further. This work is described in a report specifically about the modelling [5]. The new model was used in an appraisal of anakinra [6, 7].

The most important change we made in the model was to define a patient’s health state in absolute terms using the score from the Stanford HAQ. This scale runs from 0 (best) to 3 (worst) in units of 0.125. In our modelling, we only allowed patients to have valid HAQ scores from this discrete scale.
In a similar fashion to the previous model, the short-term improvement on starting a treatment was modelled as a fixed improvement, which was now taken as a fixed decrease in HAQ, subject to the restriction that HAQ could not go below zero. A patient’s HAQ was allowed to deteriorate while on treatment. The rate of deterioration could vary by treatment, and the possibility that a patient’s HAQ may remain constant while on a biologic treatment was modelled pragmatically as an extremely long mean time to a deterioration in HAQ. The assumption that HAQ would deteriorate on quitting treatment by an amount equal to the improvement on starting a treatment was maintained, but subject to the ceiling effect that HAQ could not worsen beyond its maximum value of 3. Figure 4 shows a possible way in which a patient’s HAQ could change over time. Figure 5 shows the full cycle of changes undergone by a patient in the Birmingham Rheumatoid Arthritis Model (BRAM).

Initial patient characteristics could now be defined as age, gender and HAQ score, with the overall distribution...
appropriate for a population starting DMARDs. In the BRAM, any patient not reaching the point of divergence between strategies was discarded from the analysis and not included in the total population count.

Instead of sampling a patient’s remaining lifetime as an initial characteristic of the patient, risk of mortality was allowed to vary with changing HAQ score. Costs were accumulated on the same basis as before. QALYs were now calculated by using the results of a regression analysis of quality of life as a function of HAQ: this analysis indicated that a linear function was appropriate, based on data. Costs and QALYs were now only counted from the point of divergence between strategies, and again were discounted to that point at 6 and 1.5%, respectively.

In the work leading to the methodological report [5], we included coding in the model to consider joint replacement and hospitalization. However, we never found sufficiently reliable data to model these explicitly in any actual appraisals, and instead used a variable HAQ-related cost as a proxy.

A further new feature in the BRAM was that instead of the underlying DMARD sequence being fixed in the model coding, it was introduced to the model through a data file, thus allowing alternative sequences to be compared more easily than before. The increased complexity of the model entailed a change in software and the BRAM was coded in Borland Delphi, based on underlying code in the fast programming language Pascal.

**The BRAM (second version)**

The next appraisal for which the BRAM was used was the review of etanercept and infliximab, to which adalimumab was now added [8, 9]. A limitation of the previous version was the use of a fixed improvement in HAQ when starting a new treatment. In the second version, we addressed this problem by allowing the HAQ improvement on starting a treatment to be sampled separately for each individual. For each individual, the improvement was modelled by sampling from a beta distribution and multiplying the off-treatment HAQ by the sampled value, and then rounding to the nearest valid HAQ score. This approach allows for the fact that a patient with a higher (worse) HAQ has greater room for improvement, and avoids an artificial division into responders and non-responders.

We also included two stages of early withdrawal from treatment. In practice, we set these at 6 and 24 weeks, but these values were introduced into the model through data files, and could have been set to any desired values. The first stage of early withdrawal was taken to relate to immediate toxicity of the treatment, while the second stage could be either for lack of effectiveness or for toxicity. The proportions of patients withdrawing for each reason was based on available data: those withdrawing for loss of effectiveness were taken as those with the least improvement in HAQ on starting treatment, while toxicity was taken to be independent of HAQ improvement.

The NICE guidance [9] included various conditions under which biologic treatments should be discontinued. Such stopping rules can be modelled in one of two ways, explicitly or implicitly. Explicit modelling of stopping rules requires additional parameters in the model and sampling each patient’s progress against the relevant stopping rule. For realistic modelling, account must be taken of the possibility that patients may also quit treatment for other reasons. Implicit modelling of stopping rules is inherently simpler as it just requires the proportion of patients...
Fig. 5 Sequence of changes in the BRAM. Figure has been reproduced with permission from 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, with permission from NIHR Evaluation, Trials and Studies Coordinating Centre, Health Technology Assessment (NETSCC, HTA).

Needs for future modelling

The assumption that total cost of a treatment can be approximated by a fixed start-up cost followed by an annual cost over a survival time taken from a continuous distribution is based on frequent treatment. This approximation is certainly reasonable for all conventional DMARDs, and for adalimumab, etanercept and abatacept. It may be a reasonable approximation for infliximab, where treatments are given every 8 weeks, but becomes seriously problematic for rituximab, where courses of treatment are separated by at least 6 months. In future modelling, it will be appropriate to make each new prescription a single event in the model. Use of a continuous time individual sampling model can accommodate a variable time to retreatment for rituximab.

All versions of the BRAM developed so far use HAQ as the sole measure of a patient’s health state. It would be desirable to include a more detailed description of patient health, possibly using aspects of health not captured in the HAQ, and possibly by considering individual components of HAQ. This approach would allow a more realistic quality of life equation than has been used up to now. It would, however, require some means of estimating the way in which all modelled patient characteristics vary over time on all treatments.

Further possibilities for development of the model include explicit modelling of adverse events, joint replacement and hospitalization. These features have been included implicitly in the modelling by increasing the annual costs of treatments, and allowing the costs to vary by HAQ score. Explicit inclusion would be preferable if it could be done without introducing excessive complication into the model.

Discussion

Decisions on treatments for RA affect a large number of people, including RA patients, other patients who may not be treated if money is spent on RA patients, rheumatologists, clinicians treating other patients, policy makers affected by policy decisions and administrators charged
with implementing policies once they have been made. It is in the interest of all of these people that such decisions are made using an appropriate synthesis of what is, and is not, known about the short- and long-term effects of treatment. This synthesis is best achieved through the use of a model with appropriate sensitivity analysis.

Any model is necessarily a simplification of reality, and is constrained by resources of data and the need for timely reporting of results. This article has illustrated the way that it is possible to develop a model over time. Table 1 summarizes the various changes that have been made. As the model is used for a range of different decisions, further issues may become pertinent. In the case of the BRAM, these have included sequential use of biologic therapies and the timing of rituximab treatments.

**Rheumatology key message**

- A continuous time individual sampling model is appropriate for modelling RA treatments.

### Acknowledgements

The BRAM has been developed by members of the West Midlands Health Technology Assessment Collaboration, named as authors of the various references. This article was written by its author alone. The development of the BRAM has been funded through a contract between the West Midlands Health Technology Assessment Collaboration and the National Co-ordinating Centre for Health Technology Assessment. No specific funding was obtained for the present article.

**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 author has declared no conflicts of interest.

### References


### Table 1

<table>
<thead>
<tr>
<th>Feature</th>
<th>BPM</th>
<th>BRAM 1</th>
<th>BRAM 2</th>
<th>BRAM 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment sequence</td>
<td>Fixed</td>
<td>Variable</td>
<td>Variable</td>
<td>Variable</td>
</tr>
<tr>
<td>Patient health state</td>
<td>QoL</td>
<td>HAQ</td>
<td>HAQ</td>
<td>HAQ</td>
</tr>
<tr>
<td>HAQ to QoL conversion</td>
<td>NA</td>
<td>Linear</td>
<td>Linear</td>
<td>Quadratic</td>
</tr>
<tr>
<td>Short-term effect of starting treatment</td>
<td>Additive fixed</td>
<td>Additive fixed</td>
<td>Multiplicative variable</td>
<td>Multiplicative variable</td>
</tr>
<tr>
<td>Mortality dependent on HAQ</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Short-term treatment quitting</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Probabilistic sensitivity analysis</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

BPM: Birmingham Preliminary Model; BRAM 1: BRAM Version 1, etc.; QoL: quality of life; HAQ = Stanford HAQ; NA: not applicable.