Rituximab for the treatment of rheumatoid arthritis

Premeeting briefing

This briefing presents major issues arising from the manufacturer’s submission (MS), evidence review group (ERG) report and personal statements made by nominated clinical specialists and patient experts. Please note that although condensed summary information is included for ease of reference, this briefing should be read in conjunction with the full supporting documents.

The manufacturer was asked to clarify the following: effectiveness data (literature searches, inclusion/exclusion criteria in the pivotal trial, differing population sizes in the different analyses, missing data, formatting errors and methodology adopted for the indirect comparisons), economic analyses (early withdrawals, inter-treatment interval, half cycle correction and inpatient resource use) and economic structure (the method of randomisation and probabilistic sensitivity analysis [PSA]).

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACR</td>
<td>American College of Rheumatology</td>
</tr>
<tr>
<td>DAS</td>
<td>disease activity score</td>
</tr>
<tr>
<td>DMARD</td>
<td>disease-modifying anti-rheumatic drug</td>
</tr>
<tr>
<td>EMEA</td>
<td>European Medicines Evaluation Agency</td>
</tr>
<tr>
<td>ERG</td>
<td>evidence review group</td>
</tr>
<tr>
<td>EULAR</td>
<td>European League Against Rheumatism</td>
</tr>
<tr>
<td>FACIT-F</td>
<td>Function Assessment of Chronic Illness Therapy Fatigue</td>
</tr>
<tr>
<td>HAQ</td>
<td>Health Assessment Questionnaire</td>
</tr>
<tr>
<td>HAQ-DI</td>
<td>Health Assessment Questionnaire Disability Index</td>
</tr>
<tr>
<td>ICER</td>
<td>incremental cost-effectiveness ratio</td>
</tr>
<tr>
<td>MS</td>
<td>manufacturer’s submission</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Clinical Excellence</td>
</tr>
<tr>
<td>PSA</td>
<td>probabilistic sensitivity analysis</td>
</tr>
<tr>
<td>QALY</td>
<td>quality-adjusted life year</td>
</tr>
<tr>
<td>RA</td>
<td>rheumatoid arthritis</td>
</tr>
<tr>
<td>RCT</td>
<td>randomised controlled trial</td>
</tr>
<tr>
<td>SF-36</td>
<td>Short Form 36</td>
</tr>
<tr>
<td>TNF-α</td>
<td>tumour necrosis factor alpha</td>
</tr>
</tbody>
</table>

Licensed indication

Rituximab (MabThera, Roche) in combination with methotrexate is indicated for the treatment of adults with severe active rheumatoid arthritis (RA) who have had an inadequate response to or intolerance of other disease-modifying...
anti-rheumatic drugs (DMARDs) including one or more tumour necrosis factor alpha (TNF-α) inhibitors.

Key issues for consideration

Decision problem

• The manufacturer has provided two different RA management scenarios as comparators. The second scenario concerns sequential use of TNF-α inhibitors. This is currently outside NICE guidance. Does the Committee consider it appropriate to appraise this scenario?

Clinical effectiveness

• Given that the pivotal trial does not compare rituximab with the relevant comparator as stated in the manufacturer’s decision problem, to what extent can this trial inform the decision on the relative clinical effectiveness of rituximab?
• Is it appropriate to extrapolate the results reported in the pivotal trial to the population in the decision problem? (Note that 40% of REFLEX trial patients do not match the patients in scenario one – the scenario based on current NICE guidance.)

Cost effectiveness

• Given the concerns raised by the ERG on the indirect comparison method adopted by the manufacturer, is the Committee confident that the adjusted American College of Rheumatology (ACR) responses used in the economic model are valid?
• Which of the different approaches to calculating Health Assessment Questionnaire (HAQ) progression rates for both active treatment and palliative care presented by the manufacturer and the ERG does the Committee feel is the most appropriate?
• Given the different approaches presented by the manufacturer and the ERG to calculate loss of efficacy (impact on HAQ score) during treatment, which approach does the Committee feel is the most appropriate?
CONFIDENTIAL

- Given the difference in the cost-effectiveness results presented by the manufacturer and the ERG, which set of assumptions do the Committee believe are most realistic?
- Does the Committee consider that there are any possible subgroups which would yield better economic results than the whole patient population as described in the decision problem?

1 Decision problem

1.1 Decision problem approach in the MS

<table>
<thead>
<tr>
<th>Population</th>
<th>Adults with severe active RA who have had an inadequate response to or intolerance of other DMARDs including one or more TNF-α inhibitors.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention</td>
<td>rituximab in combination with methotrexate</td>
</tr>
<tr>
<td>Comparators</td>
<td>There were two potential groups of pharmacological comparators considered.</td>
</tr>
<tr>
<td>1. Return to DMARDs</td>
<td>Possible treatment sequence following TNF-α inhibitor failure (etanercept assumed to be used as first TNF-α inhibitor):</td>
</tr>
<tr>
<td></td>
<td>1. Leflunomide</td>
</tr>
<tr>
<td></td>
<td>2. Gold</td>
</tr>
<tr>
<td></td>
<td>3. Ciclosporin</td>
</tr>
<tr>
<td></td>
<td>4. Palliative care (methotrexate)</td>
</tr>
<tr>
<td>2. TNF-α inhibitors used sequentially</td>
<td>Possible treatment sequence following TNF-α inhibitor failure (etanercept assumed to be used as first TNF-α inhibitor):</td>
</tr>
<tr>
<td></td>
<td>1. Adalimumab</td>
</tr>
<tr>
<td></td>
<td>2. Infliximab</td>
</tr>
<tr>
<td></td>
<td>3. Leflunomide</td>
</tr>
<tr>
<td></td>
<td>4. Gold</td>
</tr>
<tr>
<td></td>
<td>5. Ciclosporin</td>
</tr>
<tr>
<td></td>
<td>6. Palliative care (methotrexate)</td>
</tr>
<tr>
<td>Outcomes</td>
<td>The outcome measures considered as per the original scope were: physical function, joint damage, pain, fatigue, adverse effects of treatment, mortality and health-related quality of life.</td>
</tr>
<tr>
<td></td>
<td>Additional specific outcome measures considered were: ACR scores, disease activity score (DAS), European League Against Rheumatism (EULAR) scores, HAQ score, Function Assessment of Chronic Illness Therapy Fatigue (FACIT-F) score, Short form (SF)-36 scores and the Sharp radiographic assessment scores.</td>
</tr>
</tbody>
</table>
1.2 **ERG comments on the MS**

1.2.1 **Population**
The ERG made no comment on the appropriateness of the population.

1.2.2 **Intervention**
The ERG made no comment on the appropriateness of the intervention.

1.2.3 **Comparators**
The MS offers two different RA management strategies. The first strategy adheres to current NICE guidance, whereas the second falls outside current NICE guidance. Both scenarios begin with the assumption that initial treatment with one TNF-α inhibitor (etanercept, currently the most widely-used TNF-α inhibitor in England and Wales) has failed in all patients.

The ERG stated that given the uncertainty surrounding the many treatment pathways for people with RA, the appropriateness of the treatment pathways considered by the manufacturer in both scenarios may therefore be subject to debate within the medical community.

1.2.4 **Outcomes**
The ERG made no comment on the appropriateness of the outcomes.

1.3 **Technical leads’ comments**

1.3.1 **Population**
The characteristics of the patients in the model appear to be consistent with the licensed indication of rituximab: adults with severe active RA who have had an inadequate response to or intolerance of other DMARDs including one or more TNF-α inhibitors.

1.3.2 **Intervention**
Rituximab is appraised within the context of the licensed indication in combination with methotrexate.
1.3.3 Comparators

Palliative care, at the end of the sequences, involves treatment with an active DMARD (methotrexate). According to the MS, “the use of non-disease modifying interventions (e.g. non-steroidal anti-inflammatory drugs [including COX II selective types], and surgery) occurs primarily in tandem with disease modifying pharmacological interventions and as such Roche do not deem these to be a valid comparator when considered alone”.

1.4 Statements from professional groups and nominated experts

Many patients with RA can be treated satisfactorily with drugs such as sulphasalazine and methotrexate, but according to the statement from one of the nominated clinical experts, the recent introduction of TNF-α blocking drugs has opened up a new era in the treatment of patients with more severe RA. Unfortunately a number of patients do not respond, or respond inadequately to TNF-α blockade, and new thinking about the origins of RA has led to the view that blocking B lymphocytes in these patients may be a very useful therapeutic advance.

Another of the nominated clinical experts notes that rituximab is much easier to use than current anti-TNF therapy and current DMARDs. From the patient perspective, a course of treatment that keeps them well and acts rapidly is very advantageous. This clinical expert also notes that it is unclear how strong a response constitutes success. If patients improve considerably but do not achieve remission, should this be considered a failure or a partial response? It is also unclear how to judge when repeat treatment should be given. For example, if after 6 months patients are experiencing slightly more disease activity, should they have more treatment or should the clinician wait until they have a much higher level of activity?
2 Clinical effectiveness evidence

2.1 Clinical effectiveness in the MS

2.1.1 Direct comparison

The manufacturer reported the phase III, double blind REFLEX trial (WA17042) and open-label extension to the REFLEX trial (WA17531). Participants were randomly assigned to either rituximab and methotrexate or placebo and methotrexate. The primary efficacy end point was a response on the ACR20 at 24 weeks. Secondary end points were responses on ACR50, ACR70, DAS28, and EULAR at 24 weeks. Additional end points included scores on FACIT-F, HAQ-DI and SF-36 instruments, as well as Genant-modified Sharp radiographic scores at 24 weeks (for further details, see page 31 of the MS).

In the REFLEX trial, at 24 weeks, 51% of patients in the rituximab group reached an ACR20 response compared with 18% of patients in the placebo group. At 48 weeks, 19% of patients in the rituximab group reached an ACR20 response compared with 4% of patients in the placebo group.

At 24 and 48 weeks, all secondary efficacy outcomes, including ACR50 and ACR70 responses, were significantly different between the two groups (p < 0.002) in favour of rituximab.

Mean time to re-treatment with rituximab in the REFLEX study was 307 days. Pooled analyses showed that patients in the rituximab arm had better ACR responses from their original baseline response after their second course of rituximab.

A subgroup analysis assessed whether there were differences in time to re-treatment depending on the number of previous TNF-α inhibitor treatments. This showed that the time to re-treatment was greater for patients who had received one TNF-α inhibitor before receiving rituximab than those who had received two or more TNF-α inhibitors.
For the long-term efficacy and safety analyses, the manufacturer included data from two phase II randomised controlled trials (RCTs) (WA17043 and WA16291). These trials were excluded from the primary analysis because they included patients who had no previous exposure to a TNF-α inhibitor and who had received doses of rituximab for which no marketing authorisation has been granted.

Pooled analyses showed that acute infusion reactions were common with rituximab and constituted a significant proportion of the adverse events reported (see table 22, page 69 of the MS). The incidence of acute infusion reactions after subsequent treatment courses was generally lower than the incidence after the first infusion of rituximab.

### 2.1.2 Indirect comparison

There were no published RCTs that compared rituximab directly with any of the comparator drugs (second or third TNF-α inhibitors or leflunomide).

For the indirect comparison, DMARD (including TNF-α inhibitor) absolute adjusted ACR20 efficacy values adjusted by reference placebo (methotrexate, 26%) were as follows: rituximab (63%); leflunomide (51%); etanercept plus methotrexate (70%); infliximab plus methotrexate (59%); adalimumab (60%); abatacept (60%); gold (26%); ciclosporin (26%). See pages 62–64 of the MS for formulae and calculations of the adjusted ACR rates.

The manufacturer commented that only the trials of rituximab and abatacept were conducted in patients who matched the population of interest for the economic model. The trials relating to the other TNF-α inhibitors included patients who had experienced an inadequate response or intolerance to one or more DMARDs and therefore were from a less severe population.

### 2.2 ERG comments

#### 2.2.1 Direct comparison

The search strategy details provided by the manufacturer did not include any information on the subject index headings, the relationship between the
search terms, details of any additional searches or the data extraction strategy. The ERG was unable to reproduce the searches but was confident that all relevant trials had been identified by the manufacturer.

The ERG judged the REFLEX trial as reported in the MS to be of good methodological quality.

The results from the REFLEX trial at 24 and 48 weeks confirm that rituximab plus methotrexate was more effective than placebo plus methotrexate. Given that the patients in the trial were difficult to treat and had severe disabling disease with marked impairment of quality of life, the results of the REFLEX trial are convincing for this trial population. However, the ERG stated that this evidence cannot be used directly to answer the questions raised in the manufacturer’s analysis of the decision problem. This is because rituximab was not compared directly with a relevant comparator (e.g. leflunomide or second or third TNF-α inhibitor).

Furthermore, the ERG questioned whether the patients in the REFLEX trial were similar enough to the patients described in the rituximab management strategies in the manufacturer’s decision problem. This is because 40% of the REFLEX trial patients had received at least two TNF-α inhibitors before receiving rituximab.

The long-term efficacy data for re-treatment with rituximab from the REFLEX trial were favourable. However, results were limited by the small number of patients available for follow-up.

The available safety data from the REFLEX trial showed that patients who had received rituximab had slightly higher rates of adverse reactions than patients receiving placebo. The ERG noted that the European Medicines Evaluation Agency (EMEA) particularly emphasises the risks of infusion reactions and infection associated with rituximab. The ERG stated that this mirrors the belief that patients taking any new biological drugs require close surveillance.
2.2.2 Indirect comparisons

The only RCT evidence presented for rituximab was the comparison with placebo plus methotrexate. The ERG therefore considered it appropriate to conduct indirect comparisons to calculate absolute efficacy values for use in the economic model in order to answer the questions outlined in the manufacturer’s decision problem.

The methodology used for the indirect comparison was computationally sound. However, the ERG was not confident that the adjusted ACR scores described by the manufacturer were valid. In particular it was not clear from the submission that all the relevant clinical studies had been included. The rationale for the choice of method adopted for the indirect comparison was unclear. In addition, the method presented to adjust the ACR responses used only a single value for the reference placebo.

2.3 Technical leads’ comments

Although indirect comparisons were reported within this section of the submission, the manufacturer did not use the results as supportive data for the clinical effectiveness of rituximab.

The EMEA commented on the large differential drop-out due to lack of efficacy. They emphasise that such differences in favour of the treatment arm in trials is rather rare and, more importantly, the estimates of magnitude of effect may be exaggerated by considering all patients who dropped out as non-responders.

2.4 Statements from professional groups and nominated experts

The clinical experts stated that rituximab should be given in a secondary care setting as day case treatment. It needs supervision by a specialist nurse with experience in rheumatology or haematology. It is administered as two intravenous infusions, preferably under steroid cover to minimise any allergic reactions, with a 2-week interval between them. From the published data this
approach is often sufficient to provide a useful and sustained remission which often lasts up to a year.

According to the submission from one of the nominated clinical experts, in general, provided that rituximab is given with steroids, relatively few adverse events have been recorded. Because rituximab contains approximately 20% mouse protein, the danger of a human anti-chimeric antibody response does exist theoretically but in practice this rarely seems to be a problem.

Studies clearly show that B-cell depletion is an effective, well tolerated and easily administered form of therapy. It also has the advantage of ensuring patient compliance.

3 Cost effectiveness evidence

3.1 Cost effectiveness in the MS

The manufacturer’s economic model was a micro stimulation Markov model based on the REFLEX trial.

All patients entered the model at the start of their next treatment option after an initial TNF-α inhibitor therapy had failed. Patients may then respond within one of the three ACR response categories. The economic evaluation uses adjusted ACR response rates from the indirect comparison described in the clinical effectiveness section of the MS (see paragraph 2.1.2 above for a summary).

Patient disease progression was tracked within the model according to their HAQ. Baseline HAQ scores and changes in HAQ scores relative to ACR responses were taken from the REFLEX trial. This relationship was assumed to be equivalent across treatments. The rate of HAQ progression was also assumed to be equivalent across all treatments except for palliative care.

The manufacturer transposed the HAQ scores into quality-adjusted life years (QALYs) by using the Health Utilities Index (HUI)-3 transformation. The manufacturer noted that this approach was inconsistent with the NICE reference case as the utility scores were not derived from the EQ-5D.
instrument. The rationale provided by the manufacturer was that the HUI-3 transformation is based on the largest sample (n=2000) of patients (treated with adalimumab) and the data were collected in a clinical setting. Sensitivity analysis around the transposition method was presented.

In the original submission the results of the cost effectiveness were as follows:

<table>
<thead>
<tr>
<th>Scenario one: no sequential use of TNF-α inhibitor</th>
<th>Total QALYs</th>
<th>Difference</th>
<th>Total drug costs</th>
<th>Total costs</th>
<th>Difference</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rituximab</td>
<td>3.051</td>
<td>0.727</td>
<td>£36,003</td>
<td>£41,229</td>
<td>£10,675</td>
<td>£14,690</td>
</tr>
<tr>
<td>No rituximab</td>
<td>2.324</td>
<td></td>
<td>£24,254</td>
<td>£30,554</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario two: sequential use of TNF-α inhibitor</th>
<th>Total QALYs</th>
<th>Difference</th>
<th>Total drug costs</th>
<th>Total costs</th>
<th>Difference</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rituximab</td>
<td>3.933</td>
<td>0.526</td>
<td>£62,608</td>
<td>£66,583</td>
<td>£6,103</td>
<td>£11,601</td>
</tr>
<tr>
<td>No rituximab</td>
<td>3.407</td>
<td></td>
<td>£55,744</td>
<td>£60,480</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The manufacturer subsequently modified their model to take into account the concerns raised by the ERG with regard to methods of randomisation. This resulted in incremental cost-effectiveness ratios (ICERs) of £13,266 and £13,342 per QALY gained for scenarios one and two respectively. The manufacturer concluded that the results of the changes to the model demonstrated that the redesign of the individual sampling model with regard to its use of random numbers had little impact on the overall results.

Univariate sensitivity analyses of the original submitted model demonstrated that the model was most sensitive to variations in patient age (scenario one) and the assumed time interval between those patients who respond to treatment (scenario two). The sensitivity analysis demonstrated that the ICER of rituximab varied from £5,000 to £31,500 per QALY gained.

3.2 ERG comments

The ERG identified some aspects of the implementation of the original economic model that caused concern about its reliability for generating estimates of cost effectiveness. Two particular issues were raised with the National Institute for Health and Clinical Excellence Premeeting briefing – rheumatoid arthritis: rituximab; February 2007
NICE technical team and the manufacturer. These concerned the method of randomisation and the representation of parameter uncertainty in the PSA (see pages 54–55 of the ERG report).

The ERG undertook a validation exercise of the manufacturer’s revised model to confirm that the model logic had been consistently applied. The incremental values and cost-effectiveness ratios were sufficiently close to the submitted values to validate the model logic.

In response to the ERG’s criticism of the original PSA results and methods, the manufacturer provided an amended analysis using different measures of uncertainty. However, following an examination of the revised model, the ERG concluded that as a result of the limited scope of the revision and the continued presence of logic errors, the revised PSAs were unreliable aids to decision-making. Because of time constraints the ERG was unable to amend the manufacturer’s model (see page 76 of the ERG report for details of the limitations of the revised PSAs).

The ERG identified a number of clinical and economic problems with the revised economic model that call into question the validity of key model assumptions and the credibility of the ICERs generated. These related in particular to errors in mortality rates, the evidence for progression rates for HAQ scores, the calculation of treatment costs and errors/omissions in the estimation of inpatient costs (see pages 56-70 of the ERG report).

### 3.3 Additional analysis

The ERG undertook additional analysis by using alternative assumptions and parameters in the revised economic model (see page 77 of the ERG report). The amendments were as follows:

1. mortality probabilities were corrected – these were doubled in the original manufacturer’s model (see page 57 of ERG report)

2. the costs of drugs, administration and monitoring were amended (see section 4.3.4 of ERG report)
3. disease costs were amended (see page 71 of ERG report) to correct
   the identified error and include the costs of joint replacement

4. two HAQ progression rates (one for ‘all therapies’ and one for ‘palliative
   care’) were replaced with a single linear progression rate (see pages
   65–68 of the ERG report).

Amendments 1–3 did not substantially affect the cost effectiveness of
rituximab. However, the introduction of a single linear HAQ progression rate
resulted in an increase in the ICER. The overall effect of amendments 1–4
was to increase the manufacturer’s ICER from £14,694 to £40,873 per QALY
gained for scenario one and from £11,066 to £32,855 per QALY gained for
scenario two.

The ERG identified the following as potentially influencing model results:

1. whether the size of the benefit from each treatment is overstated,
   because loss of efficacy was assumed to be instantaneous rather than
   cumulative (see pages 71–74 of the ERG report)

2. whether the assumed mean time between the doses of rituximab was
   too conservative (see page 75 of the ERG report)

3. whether treatment sequencing in the submitted scenarios was
   suboptimal (see pages 79–80 of the ERG report).

In relation to point 1 above, the ‘worse case’ amendment for graduated loss of
efficacy (50% reduction in HAQ) increases the ICER from £40,873 to £80,198
per QALY gained for scenario one and from £32,855 to £65,558 per QALY
gained for scenario two. In relation to point 2 above, the ERG assumed a
longer mean time between doses resulting in slightly improved outcomes for
rituximab (a reduction from £40,873 to £37,002 per QALY gained for scenario
one and from £32,855 to £28,553 per QALY gained for scenario two).

In relation to point 3 above, the ERG undertook a set of tests to consider the
optimum treatment sequencing. The ERG concluded that ciclosporin and gold
could not be distinguished on either cost or outcome differences; nor could
either of the TNF-α inhibitor options (adalimumab and infliximab) be given preference over the other. However, it appears that leflunomide provides the same outcome benefits as rituximab at a reduced discounted cost per patient (£1,100 less) less, and therefore should be given before rituximab. This re-sequencing of treatments results in slightly improved ICERs for rituximab: £37,028 per QALY gained reduced from £40,873 for scenario one and £32,259 reduced from £32,855 for scenario two.

The ERG presented a two-way sensitivity analysis for different values of mean age and mean baseline HAQ score (see table 4-11, page 78 of the ERG report). Results were shown not to be very sensitive to assumptions regarding either the initial age of patients or the baseline HAQ scores, except that cost effectiveness was worsened for the very elderly.

The ERG undertook further analysis to consider whether the ICERs were different if subgroups by number of previous TNF-α inhibitors were considered separately, rather than combined as in the base case. The ERG concluded that no patient subgroups, including stratification by RF status or geographical region, could be identified which exhibit significantly better economic results than the whole cohort.

4 Authors

Nicola Hay and Ruth McAllister, on behalf of the Committee Chair (David Barnett) and the Lead Team (Darren Ashcroft and Stirling Bryan).

References

